# Chapter 9
# Modulation Classification

Jon Hamkins and Marvin K. Simon

Modulation classification is the process of deciding, based on observations of the received signal, what modulation is being used at the transmitter. It has long been an important component of noncooperative communications in which a listener desires to intercept an unknown signal from an adversary. It is also becoming increasingly important in cooperative communications, with the advent of the software-defined autonomous radio. Such a radio must configure itself, including what demodulator to use, based on the incoming signal.

In this chapter, we analyze the performance of optimum and sub-optimum modulation classifiers to discriminate $M$-ary phase-shift keying ($M$-PSK) from $M'$-ary phase-shift keying ($M'$-PSK). The measure of performance to be used is the *probability of misclassification*, i.e., the probability of deciding that $M$-PSK was transmitted when in fact $M'$-PSK was transmitted, or vice versa.

After dispensing with preliminaries in Section 9.1, we continue in Section 9.2 with a presentation of approximations to the maximum-likelihood (ML) classifier to discriminate between $M$-ary and $M'$-ary PSK transmitted on an additive white Gaussian noise (AWGN) channel and received noncoherently, partially coherently, or coherently, and when symbol timing is either known or unknown. A suboptimum classifier is shown to be ten times less complex than the ML classifier and has less than 0.1 dB of performance loss for symbol signal-to-noise ratios (SNRs) in the range $(-10,10)$ dB and any number of observed symbols. Other methods are shown to reduce complexity by a factor of 100 with less than 0.2 dB of performance loss. We also present a classifier that does not require an estimate of the symbol SNR, and in Section 9.3 we present a threshold optimization technique that improves the high-SNR performance of a previously published classifier. Complexity of the classifiers is discussed in Section 9.4. In

Section 9.5, we derive a classification error floor that exists for any classifier on any memoryless channel, even a noiseless one, by deriving a lower bound on the misclassification probability as a function of the number of observed samples. In Section 9.6, we present numerical results of each of the classifiers along with a summary comparison. In Section 9.7, we examine how symbol timing and modulation type may be jointly estimated. In Section 9.8, we show that, for the specific case of quadrature phase-shift keying (QPSK)/binary phase-shift keying (BPSK) classification, the error floor does not occur if $\pi/4$-QPSK modulation is used instead of QPSK. In Section 9.9, we follow the same ML approach as mentioned above for $M$-PSK (a special case of which is conventional QPSK) to derive the optimum and approximate classifiers for offset quadrature phase-shift keying (OQPSK) when received noncoherently over the AWGN. Examples are given for the special cases of OQPSK/BPSK and minimum-shift keying (MSK)/QPSK classifications. Finally, in Section 9.10, we discuss modulation classification in the presence of a carrier frequency offset.

## 9.1 Preliminaries

### 9.1.1 Signal Model

For ease of exposition, this chapter is limited to binary hypothesis testing in which each hypothesis occurs with equal a priori probability, although the extension to multiple hypotheses and unequal a priori probabilities can be done in the usual way [1]. Throughout, we assume $M < M'$ and each is a power of two, that the modulation type remains the same for $N$ observed received symbols, and that each point of the constellation is transmitted with equal probability. The carrier phase, modulated data, and symbol timing are assumed unknown, while $M$, $M'$, the symbol duration, signal power, noise variance, and carrier frequency[1] are assumed known.

As in Eq. (1-7), the complex baseband representation of the received $M$-PSK signal is

$$\tilde{r}(t) = \sqrt{2P_t} \sum_{n=-\infty}^{\infty} e^{j(\theta_n + \theta_c)} p(t - nT - \varepsilon T) + \tilde{n}(t) \tag{9-1}$$

where $2P_t$ is the known signal power of the complex baseband signal (the passband power is $P_t$); $\theta_n = [2q_n + (1 + (-1)^{M/2})/2]\pi/M$ is the data modulation for the $n$th $M$-PSK symbol, with independent and uniformly distributed

[1] Later on in the chapter, we shall consider modulation classification in the presence of a residual carrier frequency error that may exist after frequency correction.

$q_n \in \{0, 1, \cdots, M-1\}$; $\theta_c$ is the unknown carrier phase, uniform on $[0, 2\pi)$; $p(t)$ is a pulse shape satisfying $T^{-1}\int_0^T p^2(t)dt = 1$; $T$ is the known symbol duration; $\varepsilon$ is the unknown fractional symbol timing, uniform on [0,1); and $\tilde{n}(t)$ is a complex AWGN process with two-sided power-spectral density $N_0$ W/Hz per dimension (the passband process $n(t)$ has two-sided power spectral density (PSD) $N_0/2$ W/Hz).

The complex observables corresponding to the matched filter outputs at time instants $(n+1+\hat{\varepsilon})T,\ n = 0, 1, 2, \cdots, N-1$ are given by

$$\tilde{r}_n(\hat{\varepsilon}) = \frac{1}{T}\int_{(n+\hat{\varepsilon})T}^{(n+1+\hat{\varepsilon})T} \tilde{r}(t)\, p(t - nT - \hat{\varepsilon}T)\, dt \tag{9-2}$$

A sequence of $N$ observables corresponding to $N$ received symbols is denoted by $\tilde{\boldsymbol{r}}(\hat{\varepsilon}) = \left(\tilde{r}_1(\hat{\varepsilon}), \cdots, \tilde{r}_N(\varepsilon)\right)$. When the timing is known, the matched filter sets $\hat{\varepsilon} = \varepsilon$, which results in the observable

$$\tilde{r}_n = \tilde{r}_n(\varepsilon) = \sqrt{2P_t}e^{j(\theta_n+\theta_c)} + \tilde{n}_n, \quad n = 0, \cdots, N-1 \tag{9-3}$$

where we have dropped the symbol-timing argument. We will use the known-timing assumption throughout the remainder of the chapter, except in Section 9.7. In Eq. (9-3), $\tilde{n}_n = \tilde{n}_{n,R} + j\tilde{n}_{n,I}$ is a complex Gaussian random variable with mean zero, variance $\sigma^2 = N_0/T$ per dimension,[2] and independent components. Initially, we assume that $P_t$ and $\sigma^2$ are known at the receiver, although we will drop that assumption later. For convenience, we denote the symbol SNR as

$$\gamma_s = \frac{E_s}{N_0} = \frac{P_t}{\sigma^2} = \frac{P_t T}{N_0} \tag{9-4}$$

### 9.1.2 Conditional-Likelihood Function

The multivariate Gaussian probability density function of a complex vector $\mathbf{x}$ with mean $\mathbf{m}_x$ has the form [2, Eq. 2.99]

$$p(\mathbf{x}) = \frac{1}{\pi^N|\mathbf{C}|}\exp\left[-(\mathbf{x}-\mathbf{m}_x)^{*T}\mathbf{C}^{-1}(\mathbf{x}-\mathbf{m}_x)\right] \tag{9-5}$$

[2] See Section 11.1 for a more detailed description of the noise variance in the discrete-time model.

where $\mathbf{C}$ is the covariance matrix. The covariance matrix for $\tilde{\boldsymbol{r}}$, given $M$, $\boldsymbol{\theta} = (\theta_0, \cdots, \theta_{N-1})$, and $\theta_c$, is $\mathbf{C} = 2\sigma^2\mathbf{I}$, where $\mathbf{I}$ is the $N \times N$ identity matrix. Thus, the conditional probability density function of the complex baseband received signal $\tilde{\boldsymbol{r}}$, given $P_t$, $N_0$, $M$, $\boldsymbol{\theta}$, and $\theta_c$, is

$$p(\tilde{\boldsymbol{r}}|M,\boldsymbol{\theta},\theta_c) = \frac{1}{(2\pi\sigma^2)^N} \exp\left(-\frac{1}{2\sigma^2}\sum_{n=0}^{N-1}\left|\tilde{r}_n - \sqrt{2P_t}e^{j(\theta_n+\theta_c)}\right|^2\right) \tag{9-6}$$

This may be rewritten as

$$p(\tilde{\boldsymbol{r}}|M,\boldsymbol{\theta},\theta_c) = C\exp\left(-N\gamma_s + \mathrm{Re}\left\{\frac{\sqrt{2P_t}}{\sigma^2}e^{-j\theta_c}\sum_{n=0}^{N-1}\tilde{r}_n e^{-j\theta_n}\right\}\right) \tag{9-7}$$

where $C$ does not depend on $M$ and, therefore, drops out of the ratios we are about to form. When Eq. (9-7) is averaged over $\boldsymbol{\theta}$, we obtain what we refer to as the conditional-likelihood function (CLF), i.e., the conditional probability density function of the received vector signal $\tilde{\boldsymbol{r}}$, given $M$ and $\theta_c$. This is given by [3, Eq. B.3b]

$$\mathrm{CLF}_M(\theta_c) = C\exp\left[-N\gamma_s + \sum_{n=0}^{N-1}\ln\left(\frac{2}{M}\sum_{q=0}^{\frac{M}{2}-1}\cosh\left[x_n(q;\theta_c)\right]\right)\right] \tag{9-8}$$

where $x_n(q;\theta_c) = (\sqrt{2P_t}/\sigma^2)\mathrm{Re}\left[\tilde{r}_n e^{-j(\theta_c+[2q+(1+(-1)^{M/2})/2]\pi/M)}\right]$. We may rewrite Eq. (9-8) as

$$\mathrm{CLF}_M(\theta_c) = Ce^{-N\gamma_s}\left(\frac{2}{M}\right)^N \prod_{n=0}^{N-1}\sum_{q=0}^{\frac{M}{2}-1}\cosh\left[x_n(q;\theta_c)\right] \tag{9-9}$$

## 9.2 Modulation Classifiers

### 9.2.1 ML Classifiers

The ML modulation classifier results in the minimum probability of classification error if the modulation types occur with equal a priori probability. It can be implemented by comparing the likelihood ratio (LR) of the $N$ observed samples to a unity threshold.

**9.2.1.1 ML Noncoherent Classifier Averaging over Data, then Carrier Phase.** The CLF in Eq. (9-9) has already been averaged over the unknown data. The LR for the $M$ and $M'$ hypotheses is given by averaging over $\theta_c$, which is uniform over $[0, 2\pi)$, and forming the ratio

$$\mathrm{LR} = \frac{\mathrm{LF}_M}{\mathrm{LF}_{M'}} = \frac{E_{\theta_c}\left\{\mathrm{CLF}_M(\theta_c)\right\}}{E_{\theta_c}\left\{\mathrm{CLF}_{M'}(\theta_c)\right\}} \tag{9-10}$$

To compute the expectations in Eq. (9-10), typically hundreds of function evaluations are needed. In doing this, $\sqrt{2P_t}/\sigma^2$ must be known to compute $x_n(\cdot;\cdot)$. Note that $\gamma_s = P_t/\sigma^2$ alone is not sufficient to determine $\sqrt{2P_t}/\sigma^2$.

**9.2.1.2 ML Noncoherent Classifier Averaging over Carrier Phase, then Data.** Alternatively, we may average Eq. (9-7) over the carrier phase first. To do this we rewrite Eq. (9-7) in the form

$$p(\tilde{\boldsymbol{r}}\,|M, \boldsymbol{\theta}, \theta_c)$$

$$= C \exp\left(-N\gamma_s + \mathrm{Re}\left\{\frac{\sqrt{2P_t}}{\sigma^2}\left|\sum_{n=0}^{N-1}\tilde{r}_n e^{-j\theta_n}\right| e^{-j\left[\theta_c - \arg\left(\sum_{n=0}^{N-1}\tilde{r}_n e^{-j\theta_n}\right)\right]}\right\}\right) \tag{9-11}$$

$$= C \exp\left\{-N\gamma_s + \frac{\sqrt{2P_t}}{\sigma^2}\left|\sum_{n=0}^{N-1}\tilde{r}_n e^{-j\theta_n}\right| \cos\left[\theta_c - \arg\left(\sum_{n=0}^{N-1}\tilde{r}_n e^{-j\theta_n}\right)\right]\right\} \tag{9-12}$$

Averaging over the uniform distribution of $\theta_c$ gives

$$p\left(\tilde{\boldsymbol{r}}\,|M, \boldsymbol{\theta}\right) = \exp\left(-N\gamma_s\right) I_0\left(\frac{\sqrt{2P_t}}{\sigma^2}\left|\sum_{n=0}^{N-1}\tilde{r}_n e^{-j\theta_n}\right|\right) \tag{9-13}$$

where $I_0(x)$ is the zero-order modified Bessel function of the first kind with argument $x$. Next we average over the data phase sequence to obtain

$$\mathrm{LF}_M = p\left(\tilde{\boldsymbol{r}}\,|M\right) = \underset{\boldsymbol{\theta}}{E}\left\{C e^{-N\gamma_s} I_0\left(\frac{\sqrt{2P_t}}{\sigma^2}\left|\sum_{n=0}^{N-1}\tilde{r}_n e^{-j\theta_n}\right|\right)\right\} \tag{9-14}$$

This is feasible for small values of $N$, i.e., when the $M^N$ values of $\boldsymbol{\theta}$ are relatively manageable.

**9.2.1.3 ML Coherent/Partially Coherent Classification.** In coherent reception, the carrier phase $\theta_c$ is known, and the expectation in Eq. (9-10) degenerates to an evaluation of the CLF at a single point:

$$\mathrm{LR} = \frac{\mathrm{CLF}_M(\theta_c)}{\mathrm{CLF}_{M'}(\theta_c)}$$

This is the optimum statistic for ML classification with coherent reception.

In partially coherent reception, only partial knowledge is available about the carrier phase and thus the randomness is not completely removed. We may account for this in the classifier by using a distribution on $\theta_c$ that is different from the uniform distribution assumed in noncoherent reception. For example, a phase-locked loop that tracks a residual carrier may produce an error in its carrier phase estimate that is Tikhonov distributed [4,5]. Other randomness, due to oscillator phase noise, intersymbol interference, or phase ambiguities, for example, may also introduce nonuniform randomness to $\theta_c$.

### 9.2.2 Suboptimum Classifiers

**9.2.2.1 Coarse Integral Approximation in the LR.** One way to compute Eq. (9-10) is to accurately numerically evaluate the integrals in each of the numerator and denominator. One can see from Eq. (9-9) that the period of $\mathrm{CLF}_M(\theta)$ is $2\pi/M$. Thus, we can write the likelihood function (LF) for hypothesis $H \in \{M, M'\}$ as[3]

$$\mathrm{LF}_H = \frac{H}{2\pi}\int_0^{2\pi/H} \mathrm{CLF}_H(\theta)d\theta \cong \frac{1}{I}\sum_{i=1}^{I} \mathrm{CLF}_H\left(\frac{2i\pi}{IH}\right) \tag{9-15}$$

where the last approximation becomes an equality as $I \to \infty$ by the rectangular rule for integration. Standard integration algorithms [6] reduce computation by using a nonuniform partition of the interval, but even then, typically more than a hundred CLF evaluations are needed for an accurate integral evaluation. For the problem at hand, however, we needn't necessarily evaluate the integrals accurately. For example, for random realizations of $\tilde{\boldsymbol{r}}$, $\mathrm{CLF}_2(\theta_c)$ and $\mathrm{CLF}_4(\theta_c)$ often differ by more than an order of magnitude. If one of the CLFs is higher than the other for the entire range $\theta_c \in [0, 2\pi)$, then a single test (i.e., $I = 1$) comparing $\mathrm{CLF}_M(\theta)$ to $\mathrm{CLF}_{M'}(\theta)$ at any $\theta$ would produce the correct classification. While

[3] Note that Eq. (9-15) could have been written as $1/I\sum_{i=1}^{I} \mathrm{CLF}_H(2i\pi/(IM)])$, since $M < M'$, in which case the sample points of $\mathrm{CLF}_{M'}$ corresponding to the uniform quantization of $\theta_c$ would be the same as those of $\mathrm{CLF}_M$.

this property does not hold with high enough probability to produce near-ML performance, we may set the tunable parameter $I$ in Eq. (9-15) to values substantially smaller than required to obtain accurate integral evaluations, without significantly affecting classifier performance. This technique has also been used for MFSK classification [7].

**9.2.2.2 The Generalized Likelihood Ratio Test.** When $I = 1$ in the method above, the complexity is low, but the performance is poor. However, we can use $I = 1$ and still get near-ML performance by suitably choosing the single value of $\theta_c$ at which to evaluate $\mathrm{CLF}_M$. In particular, one can use $\tilde{\boldsymbol{r}}$ to estimate $\theta_c$ in the $M$-PSK hypothesis, specifically using its $M$th power to remove the modulation, averaging over the data symbols, and then normalizing the resulting angle:

$$\hat{\theta}_c^{(M)} = \frac{1}{M} \arg \sum_{n=0}^{N-1} \tilde{r}_n^M \tag{9-16}$$

where $\arg(z)$ denotes the angle of the complex quantity $z$. Equation (9-16) is the ML estimate of $\theta_c$ for low SNR, in the sense that it results from low-SNR approximations of the true ML estimate $\theta_{c_{\mathrm{ML}}}$ for $M$-PSK [8]. A decision based on the likelihood ratio $\mathrm{CLF}_M(\hat{\theta}_{c_{\mathrm{ML}}}^{(M)})/\mathrm{CLF}_{M'}(\hat{\theta}_{c_{\mathrm{ML}}}^{(M')})$ results from what is referred to as a *generalized likelihood ratio test (GLRT)*. A different GLRT arises when $\tilde{\boldsymbol{r}}$ is used to estimate $\boldsymbol{\theta}$ before classification [9]. We appropriately define the test based on $\mathrm{LR} = \mathrm{CLF}_M(\hat{\theta}_c^{(M)})/\mathrm{CLF}_{M'}(\hat{\theta}_c^{(M')})$ as a *quasi-generalized-likelihood ratio test (qGLRT)*, and it is a good approximation to Eq. (9-10) in the sense of being an approximation to coherent-reception classification. The carrier estimates in the two hypotheses are used with the same basic reasoning as in per-survivor processing, the method used to perform joint channel estimation and decoding. The phase ambiguities present in these estimates of $\theta_c$, namely, a $2\pi/M$ $(2\pi/M')$ rad ambiguity for the $M$-PSK ($M'$-PSK) hypothesis, do not cause a problem because, as previously mentioned, the CLFs are also periodic with period $2\pi/M(2\pi/M')$.

**9.2.2.3 Normalized Quasi-Log-Likelihood Ratio (nqLLR).** All other things being equal, we would prefer a classifier that requires the least knowledge of channel parameters. If a classifier required only $\gamma_s$, for example, it would be preferred over one requiring $P_t$ and $\sigma^2$ separately. Unfortunately, all of the classifiers discussed so far (ML, coarse integral approximate ML, qGLRT) require knowledge of both $\sqrt{2P_t}$ and $\sigma^2$ individually, and not simply their ratio.

We now discuss a classifier requiring knowledge of fewer parameters. We begin with the quasi-log-likelihood ratio (qLLR) approximation to the LLR [3,10]:[4]

$$\mathrm{qLLR} = \left| \sum_{n=0}^{N-1} \tilde{r}_n^M \right| \tag{9-17}$$

This arises from Eq. (9-10) using the low SNR approximations $\cosh(x) \cong 1 + x^2/2$ and $\ln(1+x) \cong x$, along with the approximation $\ln I_0(x) \cong x$ for large post-detection SNR. Although from its definition the qLLR itself does not require knowledge of $P_t$ and $\sigma^2$, its use in making a classification decision requires such knowledge to compute the proper threshold (see Section 9.3). This requirement is avoided by using a normalized qLLR (nqLLR) metric:

$$\mathrm{nqLLR} = \frac{\left| \sum_{n=0}^{N-1} \tilde{r}_n^M \right|}{\sum_{n=0}^{N-1} |\tilde{r}_n|^M} \tag{9-18}$$

Like the qLLR, this metric does not itself require knowledge of $P_t$, $\sigma^2$, or the ratio $\gamma_s = P_t/\sigma^2$; however, unlike the qLLR, the optimum threshold for it is also invariant to scale changes in $\tilde{\boldsymbol{r}}$—doubling $\tilde{\boldsymbol{r}}$ does not affect the nqLLR metric, for example—and it is fairly insensitive to variations of $P_t$ or $\sigma^2$ as well [11,12].

## 9.3 Threshold Optimization

### 9.3.1 Suboptimality of Previously Derived Thresholds

A critical limitation of the qLLR metric is that it does not approximate the LLR metric precisely enough for the optimum LLR threshold (zero) to be used with success. Indeed, the qLLR metric is always nonnegative, meaning that with a zero threshold (optimum for LLR), it would produce the same decision regardless of the transmitted modulation!

One can optimize the threshold for the qLLR metric itself, instead of using the optimum threshold (zero) for the LLR metric. An approximate analytic solution was given in [10], using the assumptions that the real and imaginary parts of the sum in Eq. (9-17) are jointly Gaussian and have equal variance under

---

[4] Note that the qLLR does not depend on $M'$ since, as shown in [10], for $M' > M$, the denominator of this approximate LLR is independent of $M'$.

the two hypotheses. These assumptions lead to the optimized threshold of [10, Eq. 37]:[5]

$$T = (2P_t)^{M/2} V_M I_0^{-1}\left[\exp\left(\frac{N}{2V_M}\right)\right] \tag{9-19}$$

where $V_M = \sum_{l=0}^{M}[(M!)^2\gamma_s^{-l}/(2l![(M-l)!]^2])$, and $I_0^{-1}(x)$ denotes the inverse of the function $I_0(x)$. However, the equal-variance approximation breaks down at high SNR. For example, in BPSK/QPSK classification, the variance of the real part of the sum in Eq. (9-17) is $2P_tV_2$ when BPSK is sent and $2P_t(V_2+1)$ when QPSK is sent [10, Eq. (A.12)]. At high SNR, $V_2 \cong 1/2$, and thus the variance is nearly three times higher when QPSK is sent. A similar comparison shows the variance of the imaginary part is nearly twice as high when QPSK is sent. Therefore, although the threshold in Eq. (9-19) may be near-optimum at low SNR, at high SNR it will be too low.

The further approximation $I_0(x) \cong e^x$, valid for large $x$ (high post-detection SNR), leads to the approximate threshold [10, Eq. 38]

$$T = \frac{N(2P_t)^{M/2}}{2} \tag{9-20}$$

This approximate threshold actually outperforms the "optimum" threshold in Eq. (9-19) for $\gamma_s > 4$ dB.

### 9.3.2 Empirical Threshold Optimization

The analytical derivation of appropriate thresholds for the qLLR in Eqs. (9-19) and (9-20) depended on the particular form of the metric, and involved some approximation. As an alternative, we present an empirical method to optimize the threshold that can be used for any classifier metric, including the qLLR and nqLLR metrics. Empirical threshold optimization has been briefly mentioned in $M$-PSK classification work [13] and in QAM classification [14]. However, the degree to which the analytic, approximate thresholds Eqs. (9-19) and (9-20) degrade performance has not been noted. Indeed, the fact that the "approximate" threshold Eq. (9-20) outperforms the "optimum" threshold Eq. (9-19) at high SNR is an indication that the assumptions in the derivation break down at high SNR.

---

[5] The original presentation in [10] used $\tilde{\boldsymbol{r}}/\sigma$ in place of $\tilde{\boldsymbol{r}}$ in the qLLR metric in Eq. (9-17), so that for each $n$, the noise variance of $\tilde{r}_n$ was unity in each dimension, not $\sigma^2$. To account for the difference, the thresholds we give in Eqs (9-19) and (9-20) are a factor $\sigma^M = (N_0/T)^{M/2}$ times what was in the original presentation in [10].

Empirical threshold determination consists of the following process:

(1) Generate a large number of received noisy $M$-PSK and $M'$-PSK samples according to Eq. (9-3).

(2) Group the samples into blocks of length $N$, and compute a set of sample classifier metrics (e.g., by Eq. (9-17) for the qLLR classifier).

(3) Sort the metrics for $M$-PSK and $M'$-PSK received symbols separately, in order of increasing value.

(4) For each observed metric of value $\alpha$, in increasing order:

   (a) Count the number of observed $M$-PSK metrics having value less than $\alpha$.

   (b) Count the number of observed $M'$-PSK metrics having value greater than $\alpha$.

   (c) Compute the probability of misclassification when using threshold $\alpha$, using the sum of the above counts.

(5) Report the threshold that minimizes the probability of misclassification.

This procedure may be implemented efficiently enough that a desktop machine can generate about 100 million samples and determine an empirically optimum threshold in less than an hour. The following additional points should be noted:

(1) Threshold optimization need be performed only once, offline. In a practical implementation, given $N$ and $\gamma_s$, a table look-up may be used to determine the threshold.

(2) The optimum threshold for the LR metric and its approximation in Eq. (9-15) is zero and requires no empirical optimization.

(3) The optimum threshold for the qLLR metric for BPSK/QPSK classification is relatively flat over a broad region of $\gamma_s$, being near 0.6 for $\gamma_s > -5$ dB. This is a desirable characteristic, because $\gamma_s$ may not be known exactly.

(4) The optimum threshold for the nqLLR metric is also quite flat, and ranges only from about 0.15 to 0.6 for $-10$ dB $< \gamma_s <$ 10 dB. At high SNR, it is approximately equal to the optimum qLLR/$(2NP_t)$ threshold, because the qLLR/$(2NP_t)$ metric and the nqLLR metric are nearly identical in that region. Thus, the nqLLR and qLLR classifiers have similar performance, and the nqLLR has the advantage that neither $P_t$ nor $\sigma^2$ need be known.

To see how the nqLLR classifier may be used without any knowledge of $P_t$ or $\sigma^2$, consider the problem of BPSK/QPSK modulation classification. As per Eq. (9-20), for BPSK/QPSK classification, the qLLR metric can be compared to a threshold of $NP_t$. Since $\sum_{n=0}^{N-1} |\tilde{r}_n|^2$ is a good estimate of $NP_t$, the nqLLR classifier may be compared to 0.5, and gives performance about the same as the qLLR classifier. In fact, optimum thresholds for the nqLLR classifier were found to range only between 0.15 and 0.6 for $P_t \in (0, \infty)$. This compares to optimum thresholds between 0 and $\infty$ for the qLLR classifier.

## 9.4 Complexity

### 9.4.1 ML Classifier

It is suggested in [3,10] that the ML classifier is not practical. However, if computation of $E_{\theta_c}\{\mathrm{CLF}(\theta_c)\}$ requires only a constant number of evaluations of $\mathrm{CLF}(\theta_c)$, then Eq. (9-10) can be computed in $O(N)$ time. This is because, under either hypothesis ($M$-PSK or $M'$-PSK), $\mathrm{CLF}(\theta_c)$ is a product of $N$ items, each of which can be computed in $O(1)$ time. In simulations with $N = 100$, $M = 2$, and $M' = 4$, it was found that 100 to 150 CLF evaluations were needed to compute the expectation, or about 100,000 hyperbolic cosine function evaluations overall.

Unfortunately, it is not true that a simple computation of $E_{\theta_c}\{\mathrm{CLF}(\theta_c)\}$ can be computed in $O(N)$ time, because the $\mathrm{CLF}(\theta_c)$ has first derivatives that are exponential in $N$, and thus the numerical computation of the integral requires finer partitions for higher values of $N$. This exponential relationship can be seen as follows. By elementary calculus, the first derivative of $\mathrm{CLF}_2(\theta_c)$ is

$$\mathrm{CLF}_2'(\theta_c) = Ce^{-N\gamma_s} \sum_{l=0}^{N-1} \left\{ \sinh\left[\frac{\sqrt{2P_t}}{\sigma^2}\mathrm{Re}\left(\tilde{r}_l e^{-j\theta_c}\right)\right] \frac{\sqrt{2P_t}}{\sigma^2}\mathrm{Im}\left(\tilde{r}_l e^{-j\theta_c}\right) \right.$$

$$\left. \times \prod_{n=0, n\neq l}^{N-1} \cosh\left[\frac{\sqrt{2P_t}}{\sigma^2}\mathrm{Re}\left(\tilde{r}_n e^{-j\theta_c}\right)\right] \right\} \tag{9-21}$$

If $\theta_c$ is small, the SNR is large, and $\tilde{\boldsymbol{r}} = \sqrt{2P_t} \cdot (1, \cdots, 1)$, then keeping terms of first order in $\theta_c$ in Eq. (9-21) yields

$$\mathrm{CLF}_2'(\theta_c) \cong -2CN\gamma_s\theta_c \left(\frac{e^{\gamma_s}}{2}\right)^N \tag{9-22}$$

Since $\mathrm{CLF}'(\theta_c)$ can be exponential in $N$, no numerical integration that partitions the domain of the integral into a number of intervals independent of $N$ will produce a correct calculation as $N$ increases. That is, for higher values of $N$, a finer partition of the domain is required and thus an increasing number of evaluations of the CLF. Therefore, computing the LR in this way requires more than $O(N)$ time.

### 9.4.2 Coarse Integral Approximate ML classifier

The approximate ML metric of Eq. (9-15) may be computed in $O(IN)$ time because it is the sum of $I$ terms, each of which is a product of $N$ terms, each of which takes $O(1)$ time to compute. If the integration range is partitioned in the same way under each hypothesis, each hyperbolic cosine evaluation used in the numerator of Eq. (9-10) is also used in the denominator, and thus, only $M'/2$ hyperbolic cosine evaluations are needed for each $(i, n)$ pair. Also, we may write

$$x_n\left(q; \frac{2i\pi}{IM}\right) = \frac{\sqrt{2P_t}}{\sigma^2}\left[\mathrm{Re}[\tilde{r}_n]\cos\big(\alpha(q,i)\big) + \mathrm{Im}[r_n]\sin\big(\alpha(q,i)\big)\right]$$

where $\alpha(q,i) = 2\pi(i/I + q)/M$. The trigonometric terms for all $\alpha(\cdot,\cdot)$ may be precomputed and stored in a table, since they do not depend on $\tilde{\boldsymbol{r}}$. Thus, only addition, multiplication, and $M'/2$ hyperbolic cosine evaluations are required for each $(i, n)$ pair.

### 9.4.3 qGLRT Classifier

The qGLRT estimator/classifier uses $I = 1$, which offers a speed advantage by a factor of 12 over the $I = 12$ case discussed above. It also requires the two phase estimates, however. These take $O(N)$ to compute and were observed to roughly cut the speed gain in half, i.e., 6 times faster than the $I = 12$ case.

### 9.4.4 qLLR and nqLLR

The qLLR and nqLLR classifiers also have $O(N)$ complexity, the lowest complexity of those considered in this chapter. They do not require any exponential, logarithmic, and trigonometric function evaluations—only multiplication, division, and magnitude operations. As a result, they were observed to offer speed increases of two orders of magnitude over ML classification.

The qLLR metric with threshold given by Eq. (9-19) or Eq. (9-20) requires knowledge of both $P_t$ and $\sigma^2$. The approximate threshold of Eq. (9-20) performs worse but does not require knowledge of $\sigma^2$. The nqLLR classifier benefits from knowledge of $\gamma_s$, but not from knowing $P_t$ and $\sigma^2$ individually. Furthermore, among all classifiers considered, it is unique in providing acceptable performance without knowledge of either $P_t$ or $\sigma^2$. If we set a fixed threshold of 0.4, performance is within approximately 1 dB of the ML classifier at $\gamma_s = 1$ and within 3.5 dB at very low $\gamma_s$.

## 9.5 Classification Error Floor

As we will see in Section 9.6, the ML classifier for $N = 10$ appears to exhibit a classification error floor at about 0.001. In this section, we analytically verify the validity of this observation by proving that an error floor of $2^{-N}$ occurs for BPSK/QPSK classification, which for $N = 10$ is about $9.8 \times 10^{-4}$. We also provide a more general error floor derivation for $M$-PSK/$M'$-PSK classification. The error floor can be explained by the probability that a randomly generated $M'$-PSK signal $N$-vector is consistent with an $M$-PSK transmission. The probability of this event can be obtained using the method of types [15]. This error floor exists for any $M$-PSK/$M'$-PSK classifier and gives an immediate lower bound on the number of observations $N$ that must be made in order to achieve a given classification error rate, even when no channel impairments are present.

**Theorem 9-1.** *If $N$ independent complex baseband symbols from an a priori equiprobable $M$-PSK or $M'$-PSK signal, $M < M'$, are observed at the output of a noncoherent noiseless channel, the minimum probability of modulation misclassification is $(M/M')^{N-1}/2$.*

**Proof.** Let $\tilde{r}_n$ be as in Eq. (9-3), with $M$-PSK and $M'$-PSK equiprobable and $\tilde{n}_n = 0$ for all $0 \le n \le N-1$. Since $\theta_c$ is uniformly distributed, $\tilde{r}_n$ is uniformly distributed on the circle of radius $\sqrt{2P_t}$, regardless of $\theta_n$. Thus, when $N = 1$ the observed signals for $M$-PSK and $M'$-PSK are identically distributed and the minimum misclassification probability is 1/2.

Now suppose $N > 1$. Let $\boldsymbol{b} = (b_0, \cdots, b_{N-1})$, where $b_n = \theta_n + \theta_c$. From $(\tilde{r}_0, \cdots, \tilde{r}_{N-1})$, we may ascertain $P_t$ and $\boldsymbol{b}$, and vice versa. Since $P_t$ does not depend on the modulation order, $\boldsymbol{b}$ is a sufficient statistic for the optimum (minimum probability of misclassification) classifier. If there exists some $n$, $0 < n \le N-1$, such that $b_n \ne b_0 \bmod 2\pi/M$, then the signal cannot be $M$-PSK, and the optimum classifier decides that $M'$-PSK was sent, with no probability of error. Otherwise, for all $n = 0, \cdots, N-1$, $b_n = b_0 \bmod 2\pi/M$. Under the $M$-PSK hypothesis, $P(b_n = b_0 \bmod 2\pi/M | M\text{-PSK}) = 1$. Under the $M'$-PSK hypothesis, for $n > 1$, $P(b_n = b_0 \bmod 2\pi/M | M'\text{-PSK}) = M/M'$, since $b_n - b_0 \bmod 2\pi/M$ takes on the $M'/M$ values in $\{0, 2\pi/M', 2 \cdot 2\pi/M', \cdots, ((M'/M) - 1) \cdot 2\pi/M')\}$ with equal probability. By the independence of the modulation symbols,

$$P(b_0 = \cdots = b_{N-1} \bmod 2\pi/M | M'\text{-PSK}) = \left(\frac{M}{M'}\right)^{N-1}$$

Thus, for all $N > 1$, the $M$-PSK-consistent event more probably arises from $M$-PSK than from $M'$-PSK, and the optimum classifier decides $M$-PSK. In summary, the optimum classifier is incorrect with probability

$$\begin{aligned} P_e &= P(M'\text{-PSK}) P_{e|M'\text{-PSK}} + P(M\text{-PSK}) P_{e|M\text{-PSK}} \\ &= \frac{1}{2} \cdot \left(\frac{M}{M'}\right)^{N-1} + \frac{1}{2} \cdot 0 = \frac{1}{2} \left(\frac{M}{M'}\right)^{N-1} \end{aligned}$$ ❐

**Corollary 9-1.** *If $N$ independent complex baseband symbols from an a priori equiprobable BPSK or QPSK signal are observed at the output of a noncoherent noiseless channel, the minimum probability of modulation misclassification is $2^{-N}$, which constitutes an error floor.*

**Corollary 9-2.** *Achieving probability of BPSK/QPSK misclassification below $10^{-6}$ requires $N \geq 20$ observed samples.*

**Proof.** $2^{-20} \cong 9.54 \times 10^{-7}$. ❐

## 9.6 Numerical Results

To illustrate the numerical performance of the proposed noncoherent classifiers, we consider the example of $M = 2$ and $M' = 4$, i.e., BPSK/QPSK classification, when the timing is known. The classifiers were each implemented in the C programming language.

The simulated performance of the ML BPSK/QPSK classifier is shown in Fig. 9-1, for various numbers of observed symbols $N$. The classification error floor of Theorem 9-1 is evident, and the high-SNR asymptote can be seen to be $2^{-N}$.

The classifier based on coarse integral approximation with $I = 12$ is shown in Fig. 9-2 for various values of $N$. The coarse integral method has performance indistinguishable from that of ML classification, and reduces the complexity (number of CLF function-evaluations) by an order of magnitude. A C implementation on a Linux desktop computer required approximately 11 milliseconds to compute the ML classification metric, and 1 millisecond to compute the coarse integral approximation metric with $I = 12$.

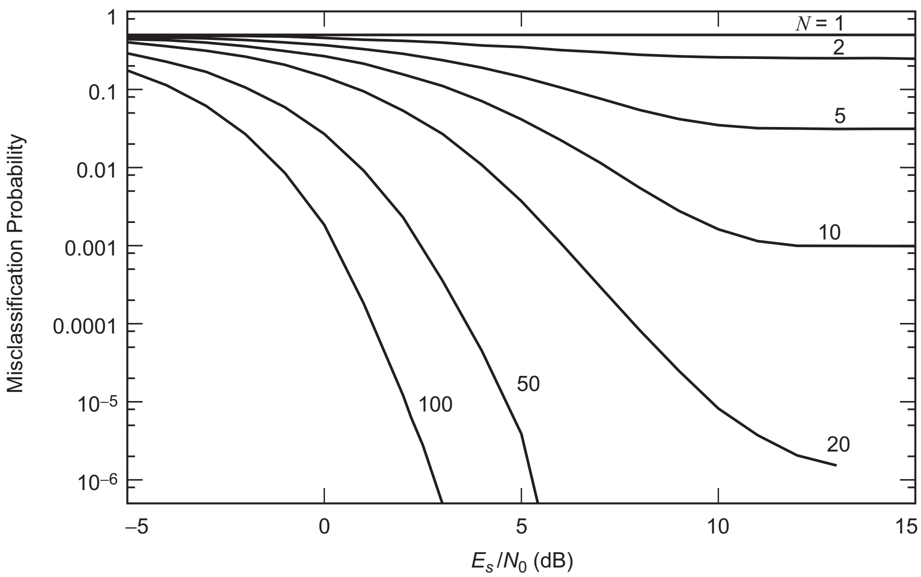


**Fig. 9-1. ML BPSK/QPSK classification performance for various $N$.**

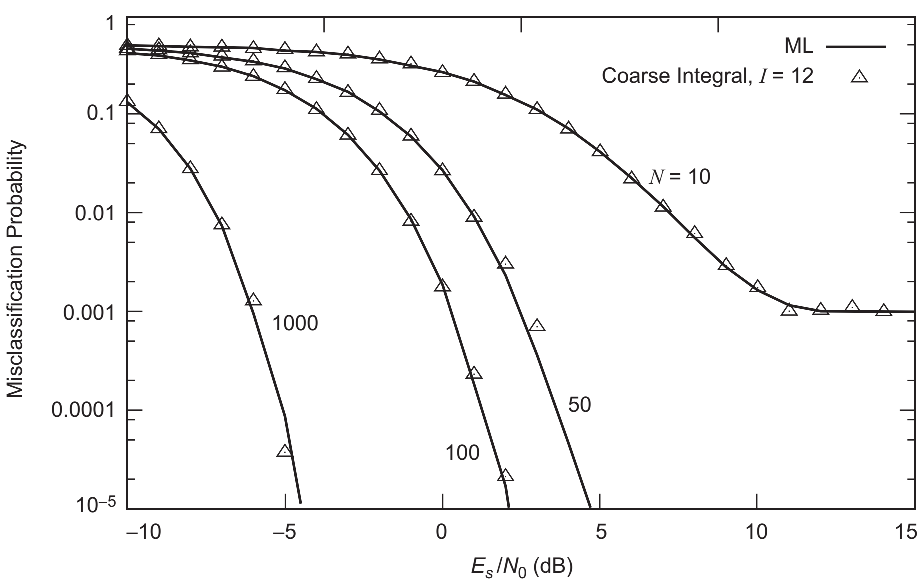


**Fig. 9-2. The performance of the ML and coarse integral classifier with $I$ = 12.**

For smaller values of $I$, performance begins to degrade. When $I = 5$, speed increased approximately by a factor of 12/5 compared to $I = 12$, and performance was similar except at high SNR, where it degraded by about 0.4 dB, as can be seen in Fig. 9-3. When $I = 1$, the classifier fails, but by choosing the single CLF evaluation point using the qGLRT, the performance improves. At high SNR, it is virtually in agreement with the ML performance, because the estimates of $\theta_c$ are quite good. At low SNR, the performance degrades slightly. The qGLRT classifier was found to be 61 times faster than the ML classifier.

For comparison, the ML coherent classifier is also illustrated in Fig. 9-3. The performance is within 0.5 of the ML noncoherent classifier over a wide range of SNR.

The effect of threshold optimization is shown in Fig. 9-4. The qLLR metric used with either of the proposed thresholds in [10] results in a classifier whose gap from ML increases with increasing $E_s/N_0$. In fact, the analytically derived threshold is outperformed by its approximation when $E_s/N_0 > 4$ dB! The empirical optimization of the threshold reduces the gap from about 2.5 dB, at $E_s/N_0 = 4$ dB, to 1 dB.

The performance of the nqLLR classifier is also shown in Fig. 9-4. Although the normalization in the nqLLR was motivated by reducing the number of parameters that need to be estimated—the nqLLR doesn't require knowledge of the noise variance—serendipitously, the normalization also has a beneficial effect in the classification performance itself. It can be seen from Fig. 9-4 that at $\gamma_s = 1$ dB, the performance of nqLLR is 0.5 dB better than qLLR, and about 0.5 from the ML limit.

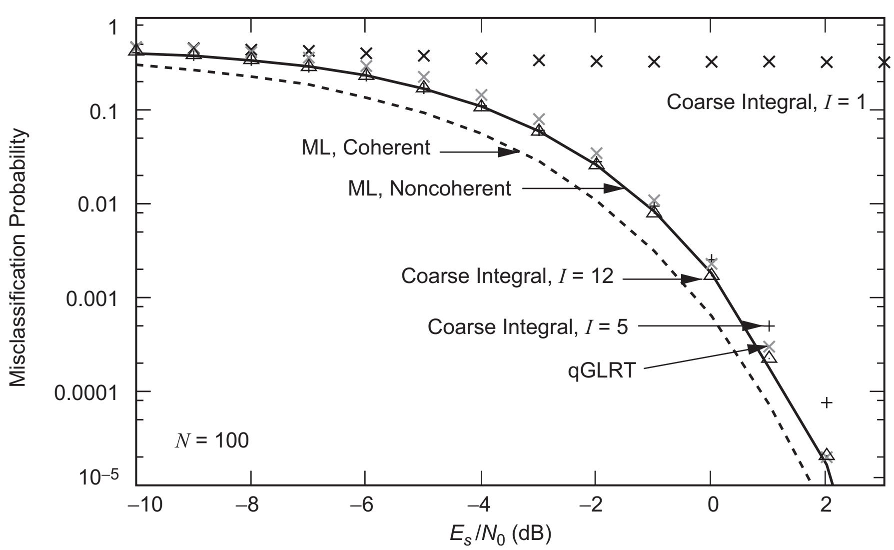


**Fig. 9-3. Noncoherent classifier performance as a function of $I$.**

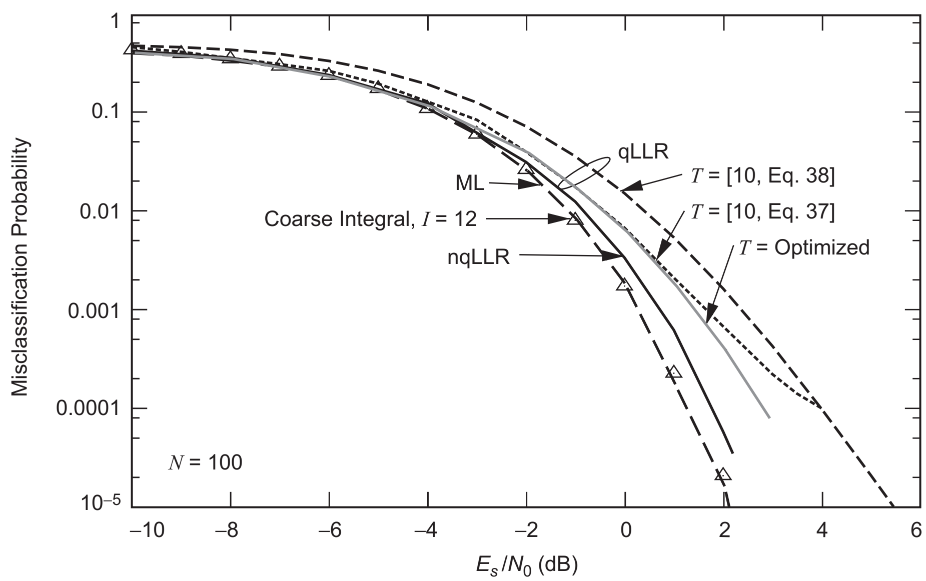


**Fig. 9-4. Comparison of ML, coarse integral classifier with $I$ = 12, nqLLR, and qLLR classifiers for various thresholds $T$.**

The values of the empirically optimized thresholds for the qLLR and nqLLR are shown in Fig. 9-5. For low SNR, the empirically optimized threshold for the qLLR metric closely matches the analytic threshold Eq. (9-19), as expected, since the analytic threshold was derived with a low-SNR approximation. At higher SNRs, the optimum threshold is near 0.6, while both Eqs. (9-19) and (9-20) converge to 0.5. The optimum nqLLR threshold also converges to about 0.6, which is also expected because the nqLLR metric converges to the qLLR metric at high SNR.

To evaluate classifier performance with imperfect symbol timing, we consider the case of a fixed fractional symbol timing offset $\tau = (\hat{\varepsilon} - \varepsilon)$, where $\tau \in (-1/2, 1/2)$. When two adjacent symbols are the same, the matched filter output does not depend on the value of $\tau$. On the other hand, when adjacent signals are antipodal, the effective SNR at the matched filter output is degraded by a factor $1 - 2|\tau|$. Thus, in the worst case, a classifier could be degraded by $10 \log_{10}(1 - 2|\tau|)$ dB in SNR. Figure 9-6 shows this upper bound for the ML BPSK/QPSK classifier when $N = 100$, as a function of $\tau$.

Table 9-1 summarizes various attributes of the classifiers. The bold entries highlight favorable performance-complexity trade-offs. If performance within 0.1 dB of the ML classifier is desired, the coarse integral method can achieve it with a speed an order of magnitude higher than the ML classifier. For operation within 0.2 dB, the qGLRT estimator/classifier or nqLLR classifier may be used,

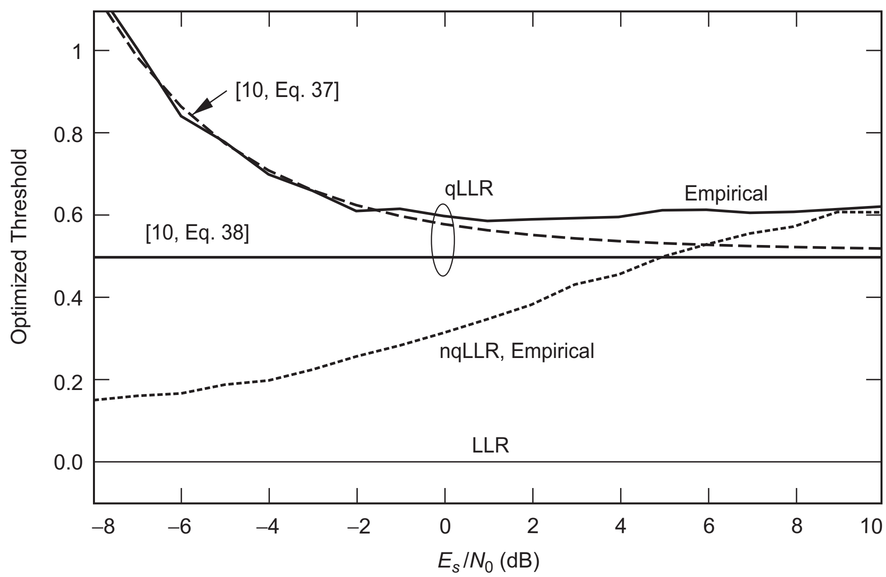


**Fig. 9-5. Dependence of empirically optimized thresholds on SNR.**

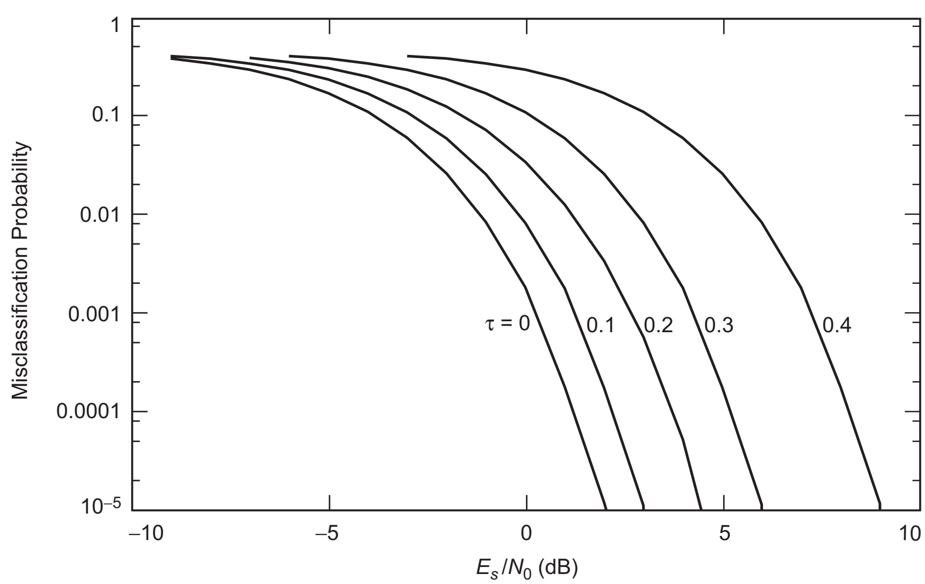


**Fig. 9-6. Upper bound on ML BPSK/QPSK classifier performance for various fixed fractional symbol timing offsets τ.**

**Table 9-1. A summary comparison of BPSK/QPSK noncoherent classifiers, $N = 100$.**

| Classifier | Equation | Threshold | Needed parameters | | dB gap to ML at $\gamma_s =$ | | | Complexity[a] | |
|---|---|---|---|---|---|---|---|---|---|
| | | | $\gamma_s$ | $P_t$ | $-5$ | $-2$ | 1 | Ops | Speed-up |
| ML | (9-10) | 1 | √ | √ | 0 | 0 | 0 | 100,000 | 1 |
| Coarse | (9-15), $I = 12$ | 1 | √ | √ | **0.0** | **0.0** | **0.1** | 7,200 | **11** |
| integral | (9-15), $I = 5$ | 1 | √ | √ | 0.0 | 0.1 | 0.4 | 3,000 | 26 |
| approx. | (9-5), $I = 1$ | 1 | √ | √ | 4.0 | 5.5 | 8.5 | 600 | 132 |
| to ML | (9-15), $I = 1$, qGLRT | 1 | √ | √ | 0.8 | 0.4 | **0.2** | 600 | **61** |
| qLLR | (9-17) | (9-19) | √ | √ | 0.4 | 0.5 | 1.1 | 0 | 165 |
| | | (9-20) | — | √ | 1.7 | 1.3 | 1.7 | 0 | 165 |
| | | Empirical | √ | √ | 0.0 | 0.5 | 1.0 | 0 | 165 |
| nqLLR | (9-18) | Empirical | — | √ | **0.0** | **0.2** | 0.5 | 0 | **99** |
| | | 0.4 | — | — | 3.5 | 2.0 | 1.0 | 0 | 99 |

[a] Ops is the number of exponential, logarithmic, trigonometric, or Bessel function evaluations required. Speed-up is the observed relative speed compared to the ML classifier.

depending on the SNR, for a speed-up factor of 61 to 99 over the ML classifier. If an SNR estimate is not available, the nqLLR metric can be used to operate within 0.5 dB of the ML classifier, at a speed-up factor of 99 over ML. The qLLR classifier is 165 times faster than ML, and can be used if losses larger than 1 dB can be tolerated.

All classifiers are subject to an error floor that is a function of the number of observed symbols, even on channels without impairments.

## 9.7 Unknown Symbol Timing

Thus far we have discussed exact and approximate likelihood-based noncoherent classifiers of $M$-PSK signals under the assumption of unknown (uniformly distributed) carrier phase, but perfectly known fractional symbol timing. In this section, we extend these classifiers to the case where the symbol timing is unknown and also uniformly distributed.

Following the same likelihood function approach as above, where we first average over the data symbol distribution, we eventually arrive at a CLF analogous to Eq. (9-8), namely,

$$\mathrm{CLF}_M\left(\theta_c,\varepsilon\right) \triangleq p\left(\tilde{\boldsymbol{r}}\,|M,\theta_c,\varepsilon\right)$$

$$= C\exp\left[-N\gamma_s + \sum_{n=0}^{N-1}\ln\left(\frac{2}{M}\sum_{q=0}^{M/2-1}\cosh\left[x_n\left(q;\theta_c,\varepsilon\right)\right]\right)\right] \tag{9-23}$$

where $\varepsilon$ denotes the unknown fractional symbol timing and

$$x_n\left(q;\theta_c,\varepsilon\right) \triangleq \frac{\sqrt{2P_t}}{\sigma^2}\mathrm{Re}\left\{\tilde{r}_n\left(\varepsilon\right)e^{-j\left(2\pi q/M+\theta_c\right)}\right\} \tag{9-24}$$

Thus, a comparison between BPSK and QPSK would be based on the LR

$$\mathrm{LR} = \frac{\underset{\theta_c,\varepsilon}{E}\left\{\mathrm{CLF}_2\left(\theta_c,\varepsilon\right)\right\}}{\underset{\theta_c,\varepsilon}{E}\left\{\mathrm{CLF}_4\left(\theta_c,\varepsilon\right)\right\}}$$

$$= \frac{\underset{\theta_c,\varepsilon}{E}\left\{\exp\left[\sum_{n=0}^{N-1}\ln\cosh\left[x_n\left(0;\theta_c,\varepsilon\right)\right]\right]\right\}}{\underset{\theta_c,\varepsilon}{E}\left\{\exp\left[\sum_{n=0}^{N-1}\ln\left(\frac{1}{2}\left[\cosh\left[x_n\left(0;\theta_c,\varepsilon\right)\right]+\cosh\left[x_n\left(1;\theta_c,\varepsilon\right)\right]\right]\right)\right]\right\}} \tag{9-25}$$

where from Eq. (9-24)[6]

$$\begin{aligned} x_n\left(0;\theta_c,\varepsilon\right) &\triangleq \frac{\sqrt{2P_t}}{\sigma^2}\mathrm{Re}\left\{\tilde{r}_n\left(\varepsilon\right)e^{-j\theta_c}\right\} \\ x_n\left(1;\theta_c,\varepsilon\right) &\triangleq \frac{\sqrt{2P_t}}{\sigma^2}\mathrm{Im}\left\{\tilde{r}_n\left(\varepsilon\right)e^{-j\theta_c}\right\} \end{aligned} \tag{9-26}$$

It has previously been shown that an efficient way of evaluating the averages over the unknown parameters, in this case $\theta_c$ and $\varepsilon$, is to apply a rectangular

[6] We slightly abuse the notation $x_n(q;\theta_c,\epsilon)$ by not explicitly listing its dependence on $M$. In this case, $x_n(0;\theta_c,\epsilon)$ for BPSK is the same as it is for QPSK, and $x_n(1;\theta_c,\epsilon)$ does not exist for BPSK. For higher orders, one needs to be more careful.

numerical integration rule. Since the CLF of Eq. (9-23) is periodic in $\theta_c$ with period $2\pi/M$, the LR of Eq. (9-25) can be expressed as

$$\mathrm{LR} = \frac{\displaystyle\lim_{I,J\to\infty} \frac{1}{I}\frac{1}{J}\sum_{i=1}^{I}\sum_{j=1}^{J} \mathrm{CLF}_2\left(\frac{i\pi}{I}, \frac{j}{J}\right)}{\displaystyle\lim_{I,J\to\infty} \frac{1}{I}\frac{1}{J}\sum_{i=1}^{I}\sum_{j=1}^{J} \mathrm{CLF}_4\left(\frac{i\pi}{2I}, \frac{j}{J}\right)} \tag{9-27}$$

To reduce the complexity of Eq. (9-27) still further, it has also been suggested one use $I = 1$ (and now also $J = 1$), i.e., evaluate the CLFs at a single value of $\theta_c$ and $\varepsilon$, where these values are obtained as functions of the same set of observables used to form the CLFs themselves. Perhaps the best set of values to use are the ML estimates of these parameters. In principle, these ML estimates should be obtained *jointly* by simultaneously maximizing the LF (or equivalently its logarithm) with respect to both $\theta_c$ and $\varepsilon$. Specifically, for the numerator of the LF we would use

$$\begin{aligned}
\hat{\theta}_{c2}, \hat{\varepsilon}_2 &= \operatorname*{argmax}_{\theta_c,\varepsilon} \sum_{n=0}^{N-1} \ln\cosh\left[x_n\left(0;\theta_c,\varepsilon\right)\right] \\
&= \operatorname*{argmax}_{\theta_c,\varepsilon} \sum_{n=0}^{N-1} \ln\cosh\left[\frac{\sqrt{2P_t}}{\sigma^2}\mathrm{Re}\left\{\tilde{r}_n\left(\varepsilon\right)e^{-j\theta_c}\right\}\right]
\end{aligned} \tag{9-28}$$

while for the denominator of the LF we would use

$$\begin{aligned}
\hat{\theta}_{c_4}, \hat{\varepsilon}_4 = \operatorname*{argmax}_{\theta_c,\varepsilon} \sum_{n=0}^{N-1} \ln\Bigg\{&\frac{1}{2}\cosh\left[\frac{\sqrt{2P_t}}{\sigma^2}\mathrm{Re}\left\{\tilde{r}_n\left(\varepsilon\right)e^{-j\theta_c}\right\}\right] \\
&+ \frac{1}{2}\cosh\left[\frac{\sqrt{2P_t}}{\sigma^2}\mathrm{Im}\left\{\tilde{r}_n\left(\varepsilon\right)e^{-j\theta_c}\right\}\right]\Bigg\}
\end{aligned} \tag{9-29}$$

Then, the low complexity LR to be used for classification would be given by

$$\mathrm{LR} = \frac{\mathrm{CLF}_2\left(\hat{\theta}_{c_2}, \hat{\varepsilon}_2\right)}{\mathrm{CLF}_4\left(\hat{\theta}_{c_4}, \hat{\varepsilon}_4\right)} \tag{9-30}$$

which when compared to a threshold again results in a GLRT. A derivation of the estimates $\hat{\theta}_{c_2}, \hat{\varepsilon}_2$ and $\hat{\theta}_{c_4}, \hat{\varepsilon}_4$ is contained in Appendix 9-A.

## 9.8 BPSK/π/4-QPSK Classification

As previously discussed in Section 9-5, for noncoherent BPSK/QPSK classification, the misclassification probability exhibits an error floor in the limit of large SNR. This occurs because, in the absence of noise, there is a finite probability that a sequence of $N$ QPSK symbols could be identical (aside from a possible fixed phase rotation over the entire sequence) to a given sequence of $N$ BPSK symbols, and thus the two modulations would be indistinguishable at the receiver, where the classification decision takes place. If instead of transmitting QPSK, one were to instead transmit $\pi/4$-QPSK [16] wherein the signal constellation is rotated back and forth by $\pi/4$ rad every other symbol, then it is no longer possible that a sequence of $N$ QPSK symbols could be "identical" to a given sequence of $N$ BPSK symbols. As such, there would no longer be an error floor in the misclassification probability performance.

In this section we investigate the ML and approximate ML algorithms appropriate to a classification decision between BPSK and $\pi/4$-QPSK.

### 9.8.1 ML Noncoherent Classifier Averaging over Data, then Carrier Phase

Analogous to the CLF for $M$-PSK in Eq. (9-9), one can derive the CLF of $\pi/4$-QPSK as

$$\begin{aligned}\mathrm{CLF}_{\pi/4-4}\left(\theta_c\right) = Ce^{-N\gamma_s}\left(\frac{1}{2}\right)^N \prod_{n=1,3,5,\cdots}^{N-1}\sum_{q=0}^{1}\cosh\left[x_n\left(q;\theta_c\right)\right] \\ \times \prod_{n=0,2,4,\cdots}^{N-2}\sum_{q=0}^{1}\cosh\left[y_n\left(q;\theta_c\right)\right]\end{aligned} \tag{9-31}$$

where

$$\begin{aligned}x_n\left(q;\theta_c\right) &= \frac{\sqrt{2P_t}}{\sigma^2}\mathrm{Re}\left\{\tilde{r}_n e^{-j(\theta_c+(2q+1)\pi/4)}\right\} \\ y_n\left(q;\theta_c\right) &= \frac{\sqrt{2P_t}}{\sigma^2}\mathrm{Re}\left\{\tilde{r}_n e^{-j(\theta_c+(2q+1)\pi/4+\pi/4)}\right\} \\ &= \frac{\sqrt{2P_t}}{\sigma^2}\mathrm{Re}\left\{\tilde{r}_n e^{-j\pi/4} e^{-j(\theta_c+(2q+1)\pi/4)}\right\}\end{aligned} \tag{9-32}$$

Since from Eq. (9-9) the CLF for BPSK would be

$$\mathrm{CLF}_2\left(\theta_c\right) = Ce^{-N\gamma_s}\prod_{n=0}^{N-1}\cosh\left[x_n\left(0;\theta_c\right)\right] \tag{9-33}$$

with $x_n\left(0;\theta_c\right) = \left(\sqrt{2P_t}/\sigma^2\right)\mathrm{Re}\left\{\tilde{r}_n e^{-j\theta_c}\right\}$, then averaging over the uniformly distributed carrier phase, the LR becomes

$$\mathrm{LR} = \frac{\underset{\theta_c}{E}\left\{\mathrm{CLF}_2\left(\theta_c\right)\right\}}{\underset{\theta_c}{E}\left\{\mathrm{CLF}_{\pi/4-4}\left(\theta_c\right)\right\}} \tag{9-34}$$

which can be evaluated numerically by the same approaches previously discussed in Section 9.2.1.

### 9.8.2 ML Noncoherent Classifier Averaging over Carrier Phase, then Data

Averaging first over the uniformly distributed carrier phase and then over the data, the unconditional LFs become [also see Eq. (9-14)]

$$\begin{aligned}
\mathrm{LF}_2 &= \underset{q}{E}\left\{Ce^{-N\gamma_s}I_0\left(\frac{\sqrt{2P_t}}{\sigma^2}\left|\sum_{n=0}^{N-1}\tilde{r}_n e^{-j\theta_n}\right|\right)\right\}, \\
&\quad \theta_1,\theta_2,\cdots,\theta_N \in (0,\pi) \\
\mathrm{LF}_{\pi/4-4} &= \underset{q}{E}\left\{Ce^{-N\gamma_s}I_0\left(\frac{\sqrt{2P_t}}{\sigma^2}\left|\sum_{n=0}^{N-1}\tilde{r}_n e^{-j\theta_n}\right|\right)\right\}, \\
&\quad \theta_1,\theta_3,\cdots,\theta_{N-1} \in \left(\frac{\pi}{4},\frac{3\pi}{4},\frac{5\pi}{4},\frac{7\pi}{4}\right), \\
&\quad \theta_2,\theta_4,\cdots,\theta_N \in \left(0,\frac{\pi}{2},\pi,\frac{3\pi}{2}\right)
\end{aligned} \tag{9-35}$$

and the corresponding LR becomes

$$\mathrm{LR} = \frac{\mathrm{LF}_2}{\mathrm{LF}_{\pi/4-4}} \tag{9-36}$$

### 9.8.3 Suboptimum Classifiers

**9.8.3.1 The GLRT.** As was the case in Section 9.2.2, finding the GLRT for the $\pi/4$-QPSK/BPSK classifier relies on finding the ML estimates of carrier phase for the two modulations under consideration. Following the approach taken in Appendix 9-A, the solution for this estimate corresponding to $\pi/4$-QPSK modulation is derived in Appendix 9-B with the result

$$\hat{\theta}_{c,\pi/4-4} = \frac{1}{4}\arg\left(\sum_{n=1}^{N}\left(\tilde{r}_n e^{-j\frac{\pi}{4}I_n}\right)^4\right) \tag{9-37}$$

where $I_n$ is the indicator variable defined by

$$I_n = \begin{cases} 0, & n \text{ odd} \\ 1, & n \text{ even} \end{cases} \tag{9-38}$$

The form of $\hat{\theta}_{c,\pi/4-4}$ is intuitively satisfying in that in the odd symbol intervals the contribution of the observable $\tilde{r}_n$ to the sum is the same as that for the QPSK ML phase estimate (i.e., $\tilde{r}_n^4$) whereas in the even symbol intervals (where the transmitted phase is shifted by $\pi/4$ rad), the observable $\tilde{r}_n$ is phase derotated by $\pi/4$ (multiplication by $e^{-j\pi/4}$) before making the same contribution to the sum.

In reality, the ML carrier phase estimate given above could be deduced immediately from the result for conventional QPSK by recognizing that $\pi/4$-QPSK can be modeled in complex baseband form as a QPSK modulator followed by multiplication by $e^{j(\pi/4)I_n}$. Thus, the corresponding ML receiver for such a modulation would be one that first undoes this alternate phase rotation, i.e., first multiplies the received signal plus noise samples by $e^{-j(\pi/4)I_n}$, and then follows that with a conventional QPSK ML receiver. (Note that multiplication by $e^{-j(\pi/4)I_n}$ does not change the statistical nature of the received noise samples.) Equivalently then, the "observables" inputs to the conventional QPSK portion of the receiver are given by $\tilde{r}_n e^{-j(\pi/4)I_n}$ and thus ML parameter estimates for $\pi/4$-QPSK are obtained from those for conventional QPSK by replacing $\tilde{r}_n$ by $\tilde{r}_n e^{-j(\pi/4)I_n}$.

For the binary case, the ML estimate remains as before, namely [see Eq. (9-16)],

$$\hat{\theta}_{c2} = \frac{1}{2}\arg\left(\sum_{n=0}^{N-1}\tilde{r}_n^2\right) \tag{9-39}$$

**9.8.3.2 The qLLR and nqLLR Metrics.** Analogous to the derivation of the qLLR metric for the BPSK/QPSK classification, one can show that the *identical* metric is appropriate for the BPSK/$\pi$/4-QPSK classification, namely,

$$\text{qLLR} = \left|\sum_{n=0}^{N-1}\tilde{r}_n^2\right| \tag{9-40}$$

and likewise for the nqLLR. Of course, the misclassification probability performance for the BPSK/$\pi$/4-QPSK classification will be different than that previously found for the BPSK/QPSK classification.

## 9.9 Noncoherent Classification of Offset Quadrature Modulations

Offset quadrature modulations are a class of modulations in which the in-phase (I) and quadrature (Q) modulations are misaligned in time with respect to one another by one-half of a symbol. Examples of such modulations are offset quadrature phase-shift keying (OQPSK) [alternatively called staggered QPSK (SQPSK)] for which the I and Q pulse streams have identical *rectangular* pulse shapes and minimum-shift keying (MSK) for which the I and Q pulse streams have identical *half-sinusoidal* pulse shapes.[7] The customary reason for using an offset form of quadrature modulation is that it reduces the maximum fluctuation of the instantaneous amplitude of the modulation from 180 deg to 90 deg since the I and Q modulations cannot change polarity at the same time instant. This has an advantage on nonlinear transmission channels where the instantaneous fluctuation of the modulation amplitude is related to the regeneration of spectral side lobes of the modulation after bandpass filtering and nonlinear amplification—the smaller the instantaneous amplitude fluctuation, the smaller the side-lobe regeneration and vice versa. On an ideal linear AWGN channel, there is no theoretical advantage of using an offset modulation relative to a conventional one; in fact, the two have identical error probability performance.

In this section, LFs for offset quadrature modulations are derived for use in the same noncoherent modulation classification applications as treated in the

[7] MSK is inherently a continuous phase frequency-shift keyed (CPFSK) modulation but can be represented as a precoded offset quadrature modulation where the precoder takes the form of a differential decoder [17, Chapter 10].

previous sections. Again, both optimum and suboptimum versions of the LFs are considered, and classification examples are presented based on discriminating OQPSK from BPSK and MSK from QPSK.

### 9.9.1 Channel Model and Conditional-Likelihood Function

For transmission of an offset quadrature modulation over an AWGN channel, the received signal can be written as

$$
\begin{aligned}
r(t) = {} & \sqrt{P_t}\left(\sum_{n=-\infty}^{\infty} a_n p(t-nT)\right)\cos(\omega_c t+\theta) \\
& - \sqrt{P_t}\left(\sum_{n=-\infty}^{\infty} b_n p(t-(n+1/2)T)\right)\sin(\omega_c t+\theta) + n(t) \qquad \text{(9-41)}
\end{aligned}
$$

where $\{a_n\}, \{b_n\}$ are independent, identically distributed (iid) binary ($\pm 1$) sequences, and as before, $T$ is the symbol time, $p(t)$ is a unit power pulse shape of duration $T$ seconds, $P_t$ is the passband signal power, and $n(t)$ is an AWGN process with single-sided power spectral density $N_0$ W/Hz. Based on the above AWGN model, then for an observation of $N$ data (symbol) intervals the CLF is given by

$$
\begin{aligned}
& p\big(r(t)|\{a_n\},\{b_n\},p(t),\theta_c\big) \\
& = \frac{1}{\sqrt{\pi N_0}}\exp\left\{-\frac{1}{N_0}\int_0^{NT}\left[r(t)-\sqrt{P_t}\left(\sum_{n=-\infty}^{\infty} a_n p(t-nT)\right)\right.\right. \\
& \qquad \left.\left.\times\cos(\omega_c t+\theta_c) + \sqrt{P_t}\left(\sum_{n=-\infty}^{\infty} b_n p(t-(n+1/2)T)\right)\sin(\omega_c t+\theta_c)\right]^2 dt\right\} \\
& = C\exp(-N\gamma)\exp\left\{\frac{2\sqrt{P_t}}{N_0}\sum_{n=0}^{N-1} a_n\int_{nT}^{(n+1)T} r(t)\,p(t-nT)\cos\omega_c t\,dt\right\} \\
& \qquad \times\exp\left\{-\frac{2\sqrt{P_t}}{N_0}\sum_{n=0}^{N-1} b_n\int_{(n+1/2)T}^{(n+3/2)T} r(t)\,p(t-(n+1/2)T)\sin\omega_c t\,dt\right\}
\end{aligned}
$$

(9-42)

where as before $\gamma_s \triangleq P_t T/N_0 = E_s/N_0$ is the symbol energy-to-noise ratio and $C$ is a constant that will be independent of the classification to be made. Transforming the received signal of Eq. (9-41) to complex baseband via $r(t) = \text{Re}\left\{\tilde{r}(t)\, e^{j\omega_c t}\right\}$ and defining the I and Q complex baseband observables (matched filter outputs) by

$$\begin{aligned}\tilde{r}_{In} &= \frac{1}{T}\int_{nT}^{(n+1)T} \tilde{r}(t)\, p(t - nT)\, dt \\ \tilde{r}_{Qn} &= \frac{1}{T}\int_{(n+1/2)T}^{(n+3/2)T} \tilde{r}(t)\, p\big(t - (n+1/2)\,T\big)\, dt\end{aligned} \tag{9-43}$$

then, letting $\sigma^2 = N_0/T$ denote the variance of these outputs, the CLF can be written in the equivalent form

$$\begin{aligned}p\left(\tilde{r}\,|\{a_n\},\{b_n\},\theta_c\right) &= C\exp(-N\gamma)\exp\left\{\frac{\sqrt{P_t}}{\sigma^2}\sum_{n=0}^{N-1} a_n \text{Re}\left\{\tilde{r}_{In}e^{-j\theta_c}\right\}\right\} \\ &\quad\times \exp\left\{\frac{\sqrt{P_t}}{\sigma^2}\sum_{n=0}^{N-1} b_n \text{Im}\left\{\tilde{r}_{Qn}e^{-j\theta_c}\right\}\right\} \\ &= C\exp\left(-N\gamma_s + \text{Re}\left\{\frac{\sqrt{P_t}}{\sigma^2}e^{-j\theta_c}\sum_{n=0}^{N-1} a_n\tilde{r}_{In}\right\}\right. \\ &\quad\left. + \text{Im}\left\{\frac{\sqrt{P_t}}{\sigma^2}e^{-j\theta_c}\sum_{n=0}^{N-1} b_n\tilde{r}_{Qn}\right\}\right)\end{aligned} \tag{9-44}$$

Note that for conventional (non-offset) QPSK we would have $\tilde{r}_{In} = \tilde{r}_{Qn} = \tilde{r}_n$, and thus letting $e^{j\theta_n} = (a_n + jb_n)/\sqrt{2}$, the CLF of Eq. (9-44) would become

$$p\left(\tilde{r}\,|\{a_n\},\{b_n\},\theta_c\right) = C\exp\left(-N\gamma_s + \text{Re}\left\{\frac{\sqrt{2P_t}}{\sigma^2}e^{-j\theta_c}\sum_{n=0}^{N-1}\tilde{r}_n e^{-j\theta_n}\right\}\right) \tag{9-45}$$

which is consistent with Eq. (9-7).

Next we average over the I and Q data streams. Before doing so, however, we first manipulate the form of Eq. (9-44) as follows:

$$p\left(\tilde{r}\left|\{a_n\},\{b_n\},\theta_c\right.\right)=C\exp\left(-N\gamma_s\right)\prod_{n=0}^{N-1}\exp\left[\mathrm{Re}\left\{\frac{\sqrt{P_t}}{\sigma^2}e^{-j\theta_c}a_n\tilde{r}_{In}\right\}\right]$$

$$\times\prod_{n=0}^{N-1}\exp\left[\mathrm{Im}\left\{\frac{\sqrt{P_t}}{\sigma^2}e^{-j\theta_c}b_n\tilde{r}_{Qn}\right\}\right]$$

$$=C\exp\left(-N\gamma_s\right)\exp\sum_{n=0}^{N-1}\ln\left(\exp\left[\mathrm{Re}\left\{\frac{\sqrt{P_t}}{\sigma^2}e^{-j\theta_c}a_n\tilde{r}_{In}\right\}\right]\right)$$

$$\times\exp\sum_{n=0}^{N-1}\ln\left(\exp\left[\mathrm{Im}\left\{\frac{\sqrt{P_t}}{\sigma^2}e^{-j\theta_c}b_n\tilde{r}_{Qn}\right\}\right]\right)\qquad\text{(9-46)}$$

Now averaging over $\{a_n\}$ and $\{b_n\}$ gives

$$p\left(\tilde{r}\left|\theta_c\right.\right)=C\exp\left(-N\gamma_s\right)\exp\sum_{n=0}^{N-1}\ln\cosh\left(\frac{\sqrt{P_t}}{\sigma^2}\mathrm{Re}\left\{\tilde{r}_{In}e^{-j\theta_c}\right\}\right)$$

$$\times\exp\sum_{n=0}^{N-1}\ln\cosh\left(\frac{\sqrt{P_t}}{\sigma^2}\mathrm{Im}\left\{\tilde{r}_{Qn}e^{-j\theta_c}\right\}\right)$$

$$=C\exp\left\{-N\gamma_s+\sum_{n=0}^{N-1}\left[\ln\cosh x_{In}\left(\theta_c\right)+\ln\cosh x_{Qn}\left(\theta_c\right)\right]\right\}\qquad\text{(9-47)}$$

where

$$\begin{aligned}x_{In}\left(\theta_c\right)&\triangleq\frac{\sqrt{P_t}}{\sigma^2}\mathrm{Re}\left\{\tilde{r}_{In}e^{-j\theta_c}\right\}\\x_{Qn}\left(\theta_c\right)&\triangleq\frac{\sqrt{P_t}}{\sigma^2}\mathrm{Im}\left\{\tilde{r}_{Qn}e^{-j\theta_c}\right\}\end{aligned}\qquad\text{(9-48)}$$

Alternatively, defining

$$
\begin{aligned}
x_n\left(0;\theta_c\right) &\triangleq x_{In}\left(\theta_c\right) - x_{Qn}\left(\theta_c\right) \\
x_n\left(1;\theta_c\right) &\triangleq x_{In}\left(\theta_c\right) + x_{Qn}\left(\theta_c\right)
\end{aligned}
\tag{9-49}
$$

and using the trigonometric identity

$$
\ln\left\{\frac{1}{2}\left[\cosh\left(X+Y\right)+\cosh\left(X-Y\right)\right]\right\} = \ln\cosh X + \ln\cosh Y \tag{9-50}
$$

the CLF of Eq. (9-47) can be manipulated into the compact form

$$
\begin{aligned}
p\left(\tilde{r}\,|\theta_c\right) &= C\exp\left\{-N\gamma_s + \sum_{n=0}^{N-1}\ln\left(\frac{1}{2}\sum_{q=0}^{1}\cosh x_n\left(q;\theta_c\right)\right)\right\} \\
&= C\exp\left(-N\gamma_s\right)\left(\frac{1}{2}\right)^N\prod_{n=0}^{N-1}\sum_{q=0}^{1}\cosh x_n\left(q;\theta_c\right) \triangleq \mathrm{CLF}_4'\left(\theta_c\right)
\end{aligned}
\tag{9-51}
$$

## 9.9.2 Classification of OQPSK versus BPSK

As an example of noncoherent classification involving an offset quadrature modulation, we consider the case of classifying OQPSK versus BPSK. As noted in previous sections, the LFs that form the LR on which the classification is based can be arrived at in two ways as follows.

**9.9.2.1 Averaging the LFs over the Data Sequences, then over the Carrier Phase.** For OQPSK, the CLF obtained by averaging the LF over the I and Q data sequences has already been determined in Eq. (9-32). The corresponding result for BPSK was previously determined as [see Eq. (9-8)]

$$
p\left(\tilde{r}\,|\theta_c\right) = C\exp\left(-N\gamma_s\right)\prod_{n=0}^{N-1}\cosh\left(\sqrt{2}x_n\left(0,\theta_c\right)\right) \triangleq \mathrm{CLF}_2\left(\theta_c\right) \tag{9-52}
$$

where for BPSK, $x_{Qn}\left(\theta_c\right) = 0$ and, hence, $x_n\left(0;\theta_c\right) = x_{In}\left(\theta_c\right)$. Thus, averaging Eqs. (9-51) and (9-52) over $\theta_c$, assumed to be uniformly distributed over the interval $\left(0, 2\pi\right)$, then the LR is computed as

$$\mathrm{LR}=\frac{\mathrm{LF}_4'}{\mathrm{LF}_2}=\frac{E_{\theta_c}\left\{\mathrm{CLF}_4'\left(\theta_c\right)\right\}}{E_{\theta_c}\left\{\mathrm{CLF}_2\left(\theta_c\right)\right\}} \tag{9-53}$$

Computing the LR from Eq. (9-53) must be performed numerically and is computationally intensive. Furthermore, in order to compute the LFs themselves, the parameters $P_t$ and $\sigma^2$ must be evaluated (to allow computation of the $x_n\left(q;\theta_c\right)$'s).

**9.9.2.2 Averaging the LFs over the Carrier Phase, then over the Data Sequences.** Suppose instead we first average the CLF of Eq. (9-44) over the carrier phase. To see how to accomplish this, we rewrite Eq. (9-44) as follows:

$$\begin{aligned}
&p\left(\tilde{r}|\{a_n\},\{b_n\},\theta_c\right)\\
&\quad= C\exp\left(-N\gamma_s+\mathrm{Re}\left\{\frac{\sqrt{P_t}}{\sigma^2}\left|\sum_{n=0}^{N-1}a_n\tilde{r}_{In}\right|e^{-j\left[\theta_c-\arg\left(\sum_{n=0}^{N-1}a_n\tilde{r}_{In}\right)\right]}\right\}\right.\\
&\qquad\left.+\mathrm{Im}\left\{\frac{\sqrt{P_t}}{\sigma^2}\left|\sum_{n=0}^{N-1}b_n\tilde{r}_{Qn}\right|e^{-j\left[\theta_c-\arg\left(\sum_{n=0}^{N-1}b_n\tilde{r}_{Qn}\right)\right]}\right\}\right)\\
&\quad= C\exp\left(-N\gamma_s+\frac{\sqrt{P_t}}{\sigma^2}\left|\sum_{n=0}^{N-1}a_n\tilde{r}_{In}\right|\cos\left[\theta_c-\arg\left(\sum_{n=0}^{N-1}a_n\tilde{r}_{In}\right)\right]\right.\\
&\qquad\left.-\frac{\sqrt{P_t}}{\sigma^2}\left|\sum_{n=0}^{N-1}b_n\tilde{r}_{Qn}\right|\sin\left[\theta_c-\arg\left(\sum_{n=0}^{N-1}b_n\tilde{r}_{Qn}\right)\right]\right)
\end{aligned} \tag{9-54}$$

Applying the trigonometric identity

$$\begin{aligned}
&|X_1|\cos(\theta_c-\phi_1)-|X_2|\sin(\theta_c-\phi_2)\\
&\quad=\sqrt{|X_1|^2+|X_2|^2+2\,|X_1|\,|X_2|\sin\left(\phi_2-\phi_1\right)}\times\cos\left(\theta_c-\eta\right)\\
&\quad=|X_1-jX_2|\cos\left(\theta_c-\eta\right);\\
&\eta=\tan^{-1}\frac{|X_1|\sin\phi_1-|X_2|\cos\phi_2}{|X_1|\cos\phi_1+|X_2|\sin\phi_2}
\end{aligned} \tag{9-55}$$

to Eq. (9-54), we get

$$
\begin{aligned}
&p\big(\tilde{r}|\{a_n\},\{b_n\},\theta_c\big) \\
&\quad = C \exp\left\{-N\gamma_s + \frac{\sqrt{P_t}}{\sigma^2}\left|\sum_{n=0}^{N-1}\left(a_n\tilde{r}_{In} - jb_n\tilde{r}_{Qn}\right)\right|\cos\left(\theta_c - \eta\right)\right\} \qquad (9\text{-}56)
\end{aligned}
$$

which when averaged over the uniform distribution of $\theta_c$ gives the desired result:

$$
p\big(\tilde{r}\,|\{a_n\}\,,\{b_n\}\big) = C\exp\left(-N\gamma_s\right) I_0\left(\frac{\sqrt{P_t}}{\sigma^2}\left|\sum_{n=0}^{N-1}\left(a_n\tilde{r}_{In} - jb_n\tilde{r}_{Qn}\right)\right|\right) \qquad (9\text{-}57)
$$

To check the consistency of this result with that for conventional QPSK, we proceed as follows. The previously derived result for the CLF of $M$-PSK is given by Eq. (9-8). For QPSK, we would have $e^{-j\theta_n} = \left(a_n - jb_n\right)/\sqrt{2}$, where $a_n$ and $b_n$ are as defined in Section 9.9.1. Thus,

$$
\left|\sum_{n=0}^{N-1}\tilde{r}_n e^{-j\theta_n}\right| = \frac{1}{\sqrt{2}}\left|\sum_{n=0}^{N-1}\tilde{r}_n\left(a_n - jb_n\right)\right| \qquad (9\text{-}58)
$$

which when substituted in Eq. (9-8) agrees with Eq. (9-57) when $\tilde{r}_{In} = \tilde{r}_{Qn} = \tilde{r}_n$, as would be the case for conventional QPSK. Finally, the LF is obtained by averaging Eq. (9-57) over the data sequences $\{a_n\}$ and $\{b_n\}$, i.e.,

$$
\mathrm{LF}_4' = E_{\{a_n\},\{b_n\}}\left\{C\exp\left(-N\gamma_s\right) I_0\left(\frac{\sqrt{P_t}}{\sigma^2}\left|\sum_{n=0}^{N-1}\left(a_n\tilde{r}_{In} - jb_n\tilde{r}_{Qn}\right)\right|\right)\right\} \qquad (9\text{-}59)
$$

which again must be done numerically and is feasible for small values of $N$.

### 9.9.3 Suboptimum (Simpler to Implement) Classifiers

In order to simplify the implementation of the ML classifiers, one must resort to approximations of the nonlinearities involved in their definitions in much the same way as was done for the conventional (non-offset) modulations. We start with the CLF averaged first over the data and then the carrier phase. Ignoring the factor $C\exp\left(-N\gamma_s\right)$ since in an LR test between two hypotheses it will cancel out, then taking the natural logarithm of Eq. (9-51) gives the log-likelihood function (LLF)

$$\text{LLF}_4' = \ln E_{\theta_c}\left\{\exp\left[\sum_{n=0}^{N-1}\ln\left(\frac{1}{2}\sum_{q=0}^{1}\cosh x_n\left(q;\theta_c\right)\right)\right]\right\} \tag{9-60}$$

Applying the approximations

$$\begin{aligned}\cosh x &\cong 1+\frac{x^2}{2}\\ \ln\left(1+x\right) &\cong x\end{aligned} \tag{9-61}$$

gives

$$\text{LLF}_4' = \ln E_{\theta_c}\left\{\exp\left[\sum_{n=0}^{N-1}\frac{1}{2}\sum_{q=0}^{1}\frac{x_n^2\left(q;\theta_c\right)}{2}\right]\right\} \tag{9-62}$$

From Eq. (9-49), we find that

$$\begin{aligned}\sum_{q=0}^{1}\frac{x_n^2\left(q;\theta_c\right)}{2} &= \frac{1}{2}\left[\left(x_{In}\left(\theta_c\right)+x_{Qn}\left(\theta_c\right)\right)^2+\left(x_{In}\left(\theta_c\right)-x_{Qn}\left(\theta_c\right)\right)^2\right]\\ &= x_{In}^2\left(\theta_c\right)+x_{Qn}^2\left(\theta_c\right)\\ &= \frac{P_t}{\sigma^4}\left[\left(\text{Re}\left\{\tilde{r}_{In}e^{-j\theta_c}\right\}\right)^2+\left(\text{Im}\left\{\tilde{r}_{Qn}e^{-j\theta_c}\right\}\right)^2\right]\end{aligned} \tag{9-63}$$

Further, using the relations

$$\begin{aligned}\left(\text{Re}\left\{\tilde{z}\right\}\right)^2 &= \frac{1}{2}\left|\tilde{z}\right|^2+\frac{1}{2}\text{Re}\left\{\tilde{z}^2\right\}\\ \left(\text{Im}\left\{\tilde{z}\right\}\right)^2 &= \frac{1}{2}\left|\tilde{z}\right|^2-\frac{1}{2}\text{Re}\left\{\tilde{z}^2\right\}\end{aligned} \tag{9-64}$$

we obtain after some simplification

$$\sum_{q=0}^{1} \frac{x_n^2\,(q;\theta_c)}{2} = \frac{P_t}{2\sigma^4}\left[|\tilde{r}_{In}|^2 + |\tilde{r}_{Qn}|^2 + \left(\mathrm{Re}\left\{\tilde{r}_{In}^2 e^{-j2\theta_c}\right\}\right)^2 - \left(\mathrm{Re}\left\{\tilde{r}_{Qn}^2 e^{-j2\theta_c}\right\}\right)^2\right] \tag{9-65}$$

which when substituted in Eq. (9-62) gives

$$\begin{aligned}\mathrm{LLF}_4' = {} & \ln E_{\theta_c}\left\{\exp\left[\sum_{n=0}^{N-1} \frac{P_t}{4\sigma^4}\left[|\tilde{r}_{In}|^2 + |\tilde{r}_{Qn}|^2 + \left(\mathrm{Re}\left\{\tilde{r}_{In}^2 e^{-j2\theta_c}\right\}\right)^2\right.\right.\right. \\ & \left.\left.\left. - \left(\mathrm{Re}\left\{\tilde{r}_{Qn}^2 e^{-j2\theta_c}\right\}\right)^2\right]\right]\right\}\end{aligned} \tag{9-66}$$

Noting that the first two terms of the summation in Eq. (9-66) do not depend on $\theta_c$, then the LLF can be simplified to

$$\begin{aligned}\mathrm{LLF}_4' = {} & \frac{P_t}{4\sigma^4}\sum_{n=0}^{N-1}\left[|\tilde{r}_{In}|^2 + |\tilde{r}_{Qn}|^2\right] \\ & + \ln E_{\theta_c}\left\{\exp\left[\frac{P_t}{4\sigma^4}\mathrm{Re}\left\{\left|\sum_{n=0}^{N-1}\tilde{r}_{In}^2\right| e^{-j\left(2\theta_c - \arg\left\{\sum_{n=0}^{N-1}\tilde{r}_{In}^2\right\}\right)}\right.\right.\right. \\ & \left.\left.\left. - \left|\sum_{n=0}^{N-1}\tilde{r}_{Qn}^2\right| e^{-j\left(2\theta_c - \arg\left\{\sum_{n=0}^{N-1}\tilde{r}_{Qn}^2\right\}\right)}\right\}\right]\right\}\end{aligned} \tag{9-67}$$

Using a relation analogous to Eq. (9-55), namely,

$$\begin{aligned}&|X_1|\cos(2\theta_c - \phi_1) - |X_2|\cos(2\theta_c - \phi_2) \\ &\quad = \sqrt{|X_1|^2 + |X_2|^2 - 2\,|X_1|\,|X_2|\cos(\phi_2 - \phi_1)} \times \cos(2\theta_c - \eta) \\ &\quad = |X_1 - X_2|\cos(2\theta_c - \eta)\,;\end{aligned} \tag{9-68}$$

$$\eta = \tan^{-1}\frac{|X_1|\sin\phi_1 - |X_2|\sin\phi_2}{|X_1|\cos\phi_1 - |X_2|\cos\phi_2}$$

the approximate LLF of Eq. (9-67) finally becomes

$$\mathrm{LLF}_4' = \frac{P_t}{4\sigma^4}\sum_{n=0}^{N-1}\left[\left|\tilde{r}_{In}\right|^2 + \left|\tilde{r}_{Qn}\right|^2\right] + \ln I_0\left(\frac{P_t}{4\sigma^4}\left|\sum_{n=0}^{N-1}\left(\tilde{r}_{In}^2 - \tilde{r}_{Qn}^2\right)\right|\right) \tag{9-69}$$

For BPSK (ignoring the same $C\exp\left(-N\gamma_s\right)$ factor), we obtain from Eq. (9-52) the LLF

$$\mathrm{LLF}_2 = \ln E_{\theta_c}\left\{\sum_{n=0}^{N-1}\ln\cosh\left(\sqrt{2}x_n\left(0,\theta_c\right)\right)\right\} \tag{9-70}$$

where again $x_n\left(0;\theta_c\right) = x_{In}\left(\theta_c\right)$ since in this case $x_{Qn}\left(\theta_c\right) = 0$. Making the same nonlinearity approximations as in Eq. (9-61), we obtain the approximate LLF

$$\begin{aligned}\mathrm{LLF}_2 =& \ln E_{\theta_c}\left\{\exp\left[\sum_{n=0}^{N-1}x_n^2\left(0;\theta_c\right)\right]\right\}\\ =& \ln E_{\theta_c}\left\{\exp\left[\frac{P_t}{\sigma^4}\sum_{n=0}^{N-1}\left(\mathrm{Re}\left\{\tilde{r}_{In}^2 e^{-j2\theta_c}\right\}\right)^2\right]\right\}\end{aligned} \tag{9-71}$$

Using Eq. (9-64), we again obtain after some simplification

$$\mathrm{LLF}_2 = \frac{P_t}{2\sigma^4}\sum_{n=0}^{N-1}\left|\tilde{r}_{In}\right|^2 + \ln I_0\left(\frac{P_t}{2\sigma^4}\left|\sum_{n=0}^{N-1}\tilde{r}_{In}^2\right|\right) \tag{9-72}$$

Finally, then qLLR$'$ is obtained as the difference of Eqs. (9-69) and (9-72), namely,

$$\begin{aligned}\mathrm{qLLR}' \triangleq \mathrm{LLF}_2 - \mathrm{LLF}_4' =& \frac{P_t}{4\sigma^4}\sum_{n=0}^{N-1}\left[\left|\tilde{r}_{In}\right|^2 - \left|\tilde{r}_{Qn}\right|^2\right] + \ln I_0\left(\frac{P_t}{2\sigma^4}\left|\sum_{n=0}^{N-1}\tilde{r}_{In}^2\right|\right)\\ &- \ln I_0\left(\frac{P_t}{4\sigma^4}\left|\sum_{n=0}^{N-1}\left(\tilde{r}_{In}^2 - \tilde{r}_{Qn}^2\right)\right|\right)\end{aligned} \tag{9-73}$$

Note once again that for a classification between BPSK and conventional QPSK where $\tilde{r}_{In} = \tilde{r}_{Qn} = \tilde{r}_n$, Eq. (9-73) would simplify to

$$\mathrm{qLLR}' \triangleq \mathrm{LLF}_2 - \mathrm{LLF}_4' = \ln I_0\left(\frac{P_t}{2\sigma^4}\left|\sum_{n=0}^{N-1}\tilde{r}_n^2\right|\right) = \ln I_0\left(\frac{\gamma}{2}\left|\sum_{n=0}^{N-1}\left(\frac{\tilde{r}_n}{\sigma}\right)^2\right|\right) \tag{9-74}$$

which is in agreement with Eq. (21) of [10].

To get to the final simplification, we now apply the approximation[8] $\ln I_0(x) \cong x \ (x \gg 1)$, resulting in

$$\mathrm{qLLR}' = \frac{P_t}{4\sigma^4}\left[\sum_{n=0}^{N-1}\left[\left|\tilde{r}_{In}\right|^2 - \left|\tilde{r}_{Qn}\right|^2\right] + 2\left|\sum_{n=0}^{N-1}\tilde{r}_{In}^2\right| - \left|\sum_{n=0}^{N-1}\left(\tilde{r}_{In}^2 - \tilde{r}_{Qn}^2\right)\right|\right] \tag{9-75}$$

Since the true LR should be compared to unity threshold, the true LLR should be compared to a zero threshold. However, as was previously discussed in Section 9.3, it is not necessarily true that, due to the nature of the approximations, the qLLR should also be compared to a zero threshold. Rather, as in the conventional QPSK/BPSK classification case, the best threshold (to minimize the probability of misclassification) should be determined by numerical means. As such, one could equivalently use

$$\mathrm{qLLR}' = \sum_{n=0}^{N-1}\left[\left|\tilde{r}_{In}\right|^2 - \left|\tilde{r}_{Qn}\right|^2\right] + 2\left|\sum_{n=0}^{N-1}\tilde{r}_{In}^2\right| - \left|\sum_{n=0}^{N-1}\left(\tilde{r}_{In}^2 - \tilde{r}_{Qn}^2\right)\right| \tag{9-76}$$

which by itself would be independent of $P_t$ and $\sigma^2$ and adjust the threshold accordingly based on knowledge of these parameters for optimum misclassification probability performance. Alternatively, as was noted for the non-offset modulation classification case in Section 9.2.2, one could avoid this threshold dependence on the signal and noise parameters by using a normalized qLLR, namely,

---

[8] Note that the argument of the Bessel function in Eq. (9-72) is large, corresponding to the post-detection SNR, i.e., after the $N$ symbols have been accumulated. This differs from the predetection SNR, $\gamma_s$, which can be assumed to be small in noisy environments.

$$\text{nqLLR}' = \frac{\sum_{n=0}^{N-1}\left[\left|\tilde{r}_{In}\right|^2 - \left|\tilde{r}_{Qn}\right|^2\right] + 2\left|\sum_{n=0}^{N-1}\tilde{r}_{In}^2\right| - \left|\sum_{n=0}^{N-1}\left(\tilde{r}_{In}^2 - \tilde{r}_{Qn}^2\right)\right|}{\sum_{n=0}^{N-1}\left[\left|\tilde{r}_{In}\right|^2 + \left|\tilde{r}_{Qn}\right|^2\right]} \tag{9-77}$$

that like the qLLR of Eq. (9-76) is itself independent of $P_t$ and $\sigma^2$, but unlike the qLLR is invariant to scale changes in $\tilde{r}$ and is fairly insensitive to variations of $P_t$ or $\sigma^2$ as well.

Finally, we conclude this section by noting that the previously discussed coarse integral approximation method applied to the individual CLFs in the LR (see Section 9.2.2) and likewise the GLRT method that replaces the integration of the CLF over $\theta_c$ by its evaluation at the ML value of $\theta_c$ can also be applied here in the offset modulation classification case.

### 9.9.4 Classification of MSK versus QPSK

Another example of classification of an offset quadrature modulation and a conventional modulation might be MSK versus QPSK. As previously noted, MSK has an offset quadrature representation in the form of a precoded OQPSK with identical half-sinusoidal pulse shapes on the I and Q channels. In order to maintain the power of the transmitted signal as $P_t$ for both modulations, the mathematical description of the received signal corresponding to transmission of MSK is in the form of Eq. (9-41), now with

$$p(t) = \sqrt{2}\sin\frac{\pi t}{T}, \quad 0 \le t \le T \tag{9-78}$$

Thus, the observables for MSK are

$$\begin{aligned}
\tilde{r}_{In} &= \sqrt{2}\,\frac{1}{T}\int_{nT}^{(n+1)T}\tilde{r}(t)\sin\frac{\pi(t-nT)}{T}dt \\
\tilde{r}_{Qn} &= \sqrt{2}\,\frac{1}{T}\int_{(n+1/2)T}^{(n+3/2)T}\tilde{r}(t)\sin\frac{\pi\left(t-(n+1/2)T\right)}{T}dt \\
&= -\sqrt{2}\,\frac{1}{T}\int_{(n+1/2)T}^{(n+3/2)T}\tilde{r}(t)\cos\frac{\pi\left(t-nT\right)}{T}dt
\end{aligned} \tag{9-79}$$

whereas the observables for QPSK are

$$\tilde{r}_{In} = \tilde{r}_{Qn} = \frac{1}{T}\int_{nT}^{(n+1)T} \tilde{r}\,(t)\,dt \tag{9-80}$$

Furthermore, from Eq. (9-48), the conditional variables needed for the CLF of MSK are

$$\begin{aligned}
x_{In}\,(\theta_c) &\triangleq \frac{\sqrt{P_t}}{\sigma^2}\,\mathrm{Re}\left\{\tilde{r}_{In}e^{-j\theta_c}\right\} \\
&= \frac{\sqrt{P_t}}{\sigma^2}\,\mathrm{Re}\left\{e^{-j\theta_c}\sqrt{2}\frac{1}{T}\int_{nT}^{(n+1)T} \tilde{r}\,(t)\sin\frac{\pi\,(t-nT)}{T}dt\right\} \\
x_{Qn}\,(\theta_c) &\triangleq \frac{\sqrt{P_t}}{\sigma^2}\,\mathrm{Im}\left\{\tilde{r}_{Qn}e^{-j\theta_c}\right\} \\
&= -\frac{\sqrt{P_t}}{\sigma^2}\,\mathrm{Im}\left\{e^{-j\theta_c}\sqrt{2}\frac{1}{T}\int_{(n+1/2)T}^{(n+3/2)T} \tilde{r}\,(t)\cos\frac{\pi\,(t-nT)}{T}dt\right\}
\end{aligned} \tag{9-81}$$

whereas those needed for the CLF of QPSK are

$$\begin{aligned}
x_{In}\,(\theta_c) &\triangleq \frac{\sqrt{P_t}}{\sigma^2}\,\mathrm{Re}\left\{\tilde{r}_{In}e^{-j\theta_c}\right\} = \frac{\sqrt{P_t}}{\sigma^2}\,\mathrm{Re}\left\{e^{-j\theta_c}\frac{1}{T}\int_{nT}^{(n+1)T} \tilde{r}\,(t)\,dt\right\} \\
x_{Qn}\,(\theta_c) &\triangleq \frac{\sqrt{P_t}}{\sigma^2}\,\mathrm{Im}\left\{\tilde{r}_{In}e^{-j\theta_c}\right\} = \frac{\sqrt{P_t}}{\sigma^2}\,\mathrm{Im}\left\{e^{-j\theta_c}\frac{1}{T}\int_{nT}^{(n+1)T} \tilde{r}\,(t)\,dt\right\}
\end{aligned} \tag{9-82}$$

Since for QPSK the CLF, namely, $\mathrm{CLF}_4\,(\theta_c)$, also has the form of Eq. (9-51), then a classification of MSK versus QPSK would be based on the LR

$$\mathrm{LR} = \frac{\mathrm{LF}_4'}{\mathrm{LF}_4} = \frac{E_{\theta_c}\left\{\mathrm{CLF}_4'\,(\theta_c)\right\}}{E_{\theta_c}\left\{\mathrm{CLF}_4\,(\theta_c)\right\}} \tag{9-83}$$

where the $x_n\,(0;\theta_c)$ and $x_n\,(1;\theta_c)$ terms in Eq. (9-49) that define the CLFs in the numerator and denominator are appropriately expressed in terms of Eqs. (9-81) and (9-82), respectively.

## 9.10 Modulation Classification in the Presence of Residual Carrier Frequency Offset

Thus far in our discussions of modulation classification, the word "noncoherent" was used to mean that the carrier phase was completely unknown, i.e., uniformly distributed in the interval $[0, 2\pi)$, but at the same time the carrier frequency was assumed to be known exactly. Here we discuss the impact on modulation classification of imperfect knowledge of the carrier frequency, i.e., the presence of a fixed residual carrier frequency offset that may exist after frequency correction. In particular, we shall point out the degrading effect of this frequency error on the behavior of the previously derived modulation classifiers and then propose an ad hoc modification of the nqLLR to cope with the problem.

When a residual radian frequency error $\Delta\omega$ is present, the received signal of Eq. (9-1) is modified to

$$\tilde{r}(t) = \sqrt{2P_t} \sum_{n=-\infty}^{\infty} e^{j(\theta_n + \theta_c + \Delta\omega t)} p(t - nT - \varepsilon T) + \tilde{n}(t) \tag{9-84}$$

or, equivalently, under the assumption of perfect symbol timing, the observables of Eq. (9-3) become

$$\tilde{r}_n = \sqrt{2P_t} \left( \frac{\sin \dfrac{\Delta\omega T}{2}}{\dfrac{\Delta\omega T}{2}} \right) e^{j(\theta_n + \theta_c + n\Delta\omega T)} + \tilde{n}_n, \quad n = 0, \cdots, N-1 \tag{9-85}$$

Based on Eq. (9-85), we see that, aside from a $\sin x/x$ amplitude factor, in effect the constellations on which the likelihood function is based are discretely rotating by $\Delta\omega T$ rad during each symbol interval of the $N$-symbol observation time. So for example, if one attempts to use the nqLLR modulation classifier of Eq. (9-18) on the received signal in Eq. (9-84), the term $\tilde{r}_n^M$ in the numerator of Eq. (9-18) still removes the $M$-PSK modulation; however, its signal component now contains the phase factor $e^{jMn\Delta\omega T}$, whose argument changes linearly throughout the summation on $n$. Thus, even in the absence of noise, one is no longer summing a set of complex observables that are aligned in phase but rather summing a set of complex observables whose phase is uniformly rotating (by increments of $M\Delta\omega T$ rad) around the circle. In fact, if the residual frequency error-symbol time product is such that $M\Delta\omega T = 2\pi/N$, then in so far as the numerator of Eq. (9-18) is concerned, the constellation will have moved through a complete revolution during the observation, thereby confusing the classification of $M$-PSK versus $M'$-PSK.

To resolve this dilemma, we need to modify the classifier in such a way as to cancel out the effect of the frequency error in the terms being summed in the numerator of the LLR so that once again they are aligned in phase. An ad hoc solution can be obtained by replacing $\tilde{r}_n^M$ with $\tilde{r}_n^M \left(\tilde{r}_{n-1}^*\right)^M$. Since the frequency error has no effect on the terms in the denominator summation of Eq. (9-18), to maintain the needed normalization for independence of the metric on $P_t$ and $\sigma^2$, we propose replacing $|\tilde{r}_n^M|$ by $|\tilde{r}_n^M \left(\tilde{r}_{n-1}^*\right)^M|$. Thus, in the presence of residual frequency error, the modification of Eq. (9-18) would become

$$\text{nqLLR} = \frac{\left|\sum_{n=1}^{N-1} \tilde{r}_n^M \left(\tilde{r}_{n-1}^*\right)^M\right|}{\sum_{n=1}^{N-1} \left|\tilde{r}_n^M \left(\tilde{r}_{n-1}^*\right)^M\right|} \tag{9-86}$$

Since, compared with Eq. (9-18), this metric applied to BPSK/QPSK classification now involves fourth-order (as opposed to second-order) signal × noise and noise × noise products, one anticipates a degradation in performance even in the case of zero residual frequency error. To compensate for this additional degradation, one would need to increase the length of the observable, i.e., increase $N$. It should also be pointed out that the optimized decision thresholds computed as in Section 9.3 will be different for the nqLLR of Eq. (9-86) than those for the nqLLR of Eq. (9-18). However, the procedure needed to perform the empirical threshold optimization would still follow the same steps as those discussed in Section 9.3.2.

## References

[1] H. V. Poor, *An Introduction to Signal Detection and Estimation*, New York: Springer-Verlag, 1988.

[2] C. W. Therrien, *Discrete Random Signal and Statistical Signal Processing*, Englewood Cliffs, New Jersey: Prentice Hall, 1992.

[3] A. Polydoros and K. Kim, "On the Detection and Classification of Quadrature Digital Modulations in Broad-Band Noise," *IEEE Transactions on Communications*, vol. 38, pp. 1199–1211, August 1990.

[4] J. H. Yuen, *Deep Space Telecommunications Systems Engineering*, New York: Plenum Press, 1983.

[5] Y. Yang and S. Soliman, "A Suboptimal Algorithm for Modulation Classification," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 33, no. 1, pp. 38–45, 1997.

[6] W. H. Press, W. T. Vetterling, S. A. Teukolsky, and B. P. Flannery, *Numerical Recipes in C*, New York: Cambridge University Press, 1992.

[7] B. Beidas and C. Weber, "Asynchronous Classification of MFSK Signals Using the Higher Order Correlation Domain," *IEEE Transactions on Communications*, vol. 46, no. 4, pp. 480–493, 1998.

[8] U. Mengali and A. N. D'Andrea, *Synchronization Techniques for Digital Receivers*, New York: Plenum, 1997.

[9] N. E. Lay and A. Polydoros, "Modulation Classification of Signals in Unknown ISI Environments," *Military Communications Conference, 1995. MILCOM'95, Conference record, IEEE*, vol. 1, San Diego, California, pp. 170–174, November 1995.

[10] C.-Y. Huang and A. Polydoros, "Likelihood Methods for MPSK Modulation Classification," *IEEE Transactions on Communications*, vol. 43, pp. 1493–1504, February/March/April 1995.

[11] A. Swami and B. Sadler, "Hierarchical Digital Modulation Classification Using Cumulants," *IEEE Transactions on Communications*, vol. 48, no. 3, pp. 416–429, 2000.

[12] O. A. Dobre, Y. Bar-Ness, and W. Su, "Higher-Order Cyclic Cumulants for High Order Modulation Classification," *Military Communications Conference, 2003. MILCOM 2003, IEEE*, vol. 1, Boston, Massachusetts, pp. 112–117, 2003.

[13] P. Panagiotou, A. Anastasopoulos, and A. Polydoros, "Likelihood Ratio Tests for Modulation Classification," *MILCOM 2000, 21st Century Military Communications Conference Proceedings*, vol. 2, Los Angeles, California, pp. 670–674, 2000.

[14] C. Long, K. Chugg, and A. Polydoros, "Further Results in Likelihood Classification of QAM Signals," *Military Communications Conference, 1994. MILCOM '94. Conference Record, 1994 IEEE*, Fort Monmouth, New Jersey, pp. 57–61, 1994.

[15] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, New York: Wiley, 1991.

[16] P. A. Baker, "Phase-Modulation Data Sets for Serial Transmission at 200 and 2400 Bits per Second, Part I," *AIEE Trans. on Commun. Electron.*, July 1962.

[17] M. K. Simon, S. M. Hinedi, and W. C. Lindsey, *Digital Communication Techniques: Signal Design and Detection*, Upper Saddle River, New Jersey: Prentice Hall, 1995.

# Appendix 9-A
# Parameter Estimation for the GLRT

To obtain the analytical form of the ML estimates of the unknown parameters in terms of the observables, one must make certain approximations to the nonlinear functions involved in Eqs. (9-28) and (9-29). In particular, for small arguments (e.g., low SNR), applying the approximations $\ln(1+x) \cong x$ and $\cosh(x) \cong 1 + x^2/2$ to Eq. (9-28) and letting $\tilde{r}_n(\varepsilon) = r_{I,n}(\varepsilon) + jr_{Q,n}(\varepsilon)$ gives

$$\begin{aligned}
\hat{\theta}_{c2}, \hat{\varepsilon}_2 &= \underset{\theta_c,\varepsilon}{\operatorname{argmax}} \sum_{n=0}^{N-1} \frac{1}{2}\left(\frac{\sqrt{2P_t}}{\sigma^2}\operatorname{Re}\left\{\tilde{r}_n(\varepsilon)\, e^{-j\theta_c}\right\}\right)^2 \\
&= \underset{\theta_c,\varepsilon}{\operatorname{argmax}} \sum_{n=0}^{N-1} \frac{1}{2}\left(\frac{\sqrt{2P_t}}{\sigma^2}\right)^2 \left(r_{In}(\varepsilon)\cos\theta_c + r_{Qn}(\varepsilon)\sin\theta_c\right)^2 \\
&= \underset{\theta_c,\varepsilon}{\operatorname{argmax}} \sum_{n=0}^{N-1} \left[\frac{1}{2}\left(r_{In}^2(\varepsilon) + r_{Qn}^2(\varepsilon)\right) + \frac{1}{2}\left(r_{In}^2(\varepsilon) - r_{Qn}^2(\varepsilon)\right)\cos 2\theta_c \right. \\
&\quad \left. + r_{In}(\varepsilon)\, r_{Qn}(\varepsilon)\sin 2\theta_c\right] \\
&= \underset{\theta_c,\varepsilon}{\operatorname{argmax}} \sum_{n=0}^{N-1} \left[\frac{1}{2}\left(r_{In}^2(\varepsilon) - r_{Qn}^2(\varepsilon)\right)\cos 2\theta_c + r_{In}(\varepsilon)\, r_{Qn}(\varepsilon)\sin 2\theta_c\right] \\
&= \underset{\theta_c,\varepsilon}{\operatorname{argmax}} \frac{1}{2}\sum_{n=0}^{N-1} \operatorname{Re}\left\{\tilde{r}_n^2(\varepsilon)\, e^{-j2\theta_c}\right\}
\end{aligned} \tag{A-1}$$

Equivalently, for a given $\varepsilon$, $\hat{\theta}_{c_2}$ is the solution of

$$\begin{aligned}
&\frac{d}{d\theta_c}\sum_{n=0}^{N-1}\left[\frac{1}{2}\left(r_{In}^2(\varepsilon)-r_{Qn}^2(\varepsilon)\right)\cos 2\theta_c + r_{In}(\varepsilon)\, r_{Qn}(\varepsilon)\sin 2\theta_c\right] \\
&\qquad = -\frac{1}{2}\sin 2\theta_c\sum_{n=0}^{N-1}\left(r_{In}^2(\varepsilon)-r_{Qn}^2(\varepsilon)\right)+\cos 2\theta_c\sum_{n=0}^{N-1} r_{In}(\varepsilon)\, r_{Qn}(\varepsilon) \\
&\qquad = \frac{1}{2}\sum_{n=0}^{N-1}\mathrm{Im}\left\{\tilde{r}_n^2(\varepsilon)\, e^{-j2\theta_c}\right\}=0
\end{aligned} \tag{A-2}$$

or

$$\hat{\theta}_{c_2}=\tan^{-1}\frac{\sum_{n=0}^{N-1} r_{In}(\varepsilon)\, r_{Qn}(\varepsilon)}{\frac{1}{2}\sum_{n=0}^{N-1}\left(r_{In}^2(\varepsilon)-r_{Qn}^2(\varepsilon)\right)} \tag{A-3}$$

which can be written in the simpler form

$$\hat{\theta}_{c_2}=\frac{1}{2}\arg\left(\sum_{n=0}^{N-1}\tilde{r}_n^2(\varepsilon)\right) \tag{A-4}$$

This estimate of $\theta_c$ is a generalization of the ML estimate when $\varepsilon = 0$ as given in Eq. (9-16), for the special case of $H = 2$.

Similarly, applying the same approximations of the nonlinearities to Eq. (9-29) gives

$$\begin{aligned}
&\hat{\theta}_{c4},\hat{\varepsilon}_4 \\
&= \operatorname*{argmax}_{\theta_c,\varepsilon}\sum_{n=0}^{N-1}\frac{1}{2}\left[\left(\frac{\sqrt{2P_t}}{\sigma^2}\mathrm{Re}\left\{\tilde{r}_n(\varepsilon)\, e^{-j\theta_c}\right\}\right)^2+\left(\frac{\sqrt{2P_t}}{\sigma^2}\mathrm{Im}\left\{\tilde{r}_n(\varepsilon)\, e^{-j\theta_c}\right\}\right)^2\right] \\
&= \operatorname*{argmax}_{\theta_c,\varepsilon}\sum_{n=0}^{N-1}\frac{1}{2}\left(\frac{\sqrt{2P_t}}{\sigma^2}\right)^2\left|\tilde{r}_n(\varepsilon)\right|^2
\end{aligned} \tag{A-5}$$

which unfortunately is independent of $\theta_c$. Thus, we see that for QPSK we must take the next term in the approximation of the hyperbolic cosine function, i.e., we should use

$$\begin{aligned} \ln(1+x) &\cong x \\ \cosh x &\cong 1 + \frac{x^2}{2} + \frac{x^4}{24} \end{aligned} \tag{A-6}$$

When this is applied to Eq. (9-29), we obtain

$$\begin{aligned}
&\hat{\theta}_{c4}, \varepsilon_4 \\
&= \underset{\theta_c, \varepsilon}{\operatorname{argmax}} \sum_{n=0}^{N-1} \frac{1}{48} \left[ \left( \frac{\sqrt{2P_t}}{\sigma^2} \operatorname{Re}\left\{ \tilde{r}_n(\varepsilon) e^{-j\theta_c} \right\} \right)^4 + \left( \frac{\sqrt{2P_t}}{\sigma^2} \operatorname{Im}\left\{ \tilde{r}_n(\varepsilon) e^{-j\theta_c} \right\} \right)^4 \right] \\
&= \underset{\theta_c, \varepsilon}{\operatorname{argmax}} \sum_{n=0}^{N-1} \frac{1}{48} \left( \frac{\sqrt{2P_t}}{\sigma^2} \right)^4 \left\{ \left[ \left( \operatorname{Re}\left\{ \tilde{r}_n(\varepsilon) e^{-j\theta_c} \right\} \right)^2 + \left( \operatorname{Im}\left\{ \tilde{r}_n(\varepsilon) e^{-j\theta_c} \right\} \right)^2 \right]^2 \right. \\
&\quad \left. -2 \left( \operatorname{Re}\left\{ \tilde{r}_n(\varepsilon) e^{-j\theta_c} \right\} \right)^2 \left( \operatorname{Im}\left\{ \tilde{r}_n(\varepsilon) e^{-j\theta_c} \right\} \right)^2 \right\} \\
&= \underset{\theta_c, \varepsilon}{\operatorname{argmin}} \sum_{n=0}^{N-1} \left( \operatorname{Re}\left\{ \tilde{r}_n(\varepsilon) e^{-j\theta_c} \right\} \right)^2 \left( \operatorname{Im}\left\{ \tilde{r}_n(\varepsilon) e^{-j\theta_c} \right\} \right)^2
\end{aligned} \tag{A-7}$$

After some manipulation it can be shown that, analogous to Eq. (A-4), the ML estimate of $\theta_c$ for a given $\varepsilon$ is given by

$$\hat{\theta}_{c_4} = \frac{1}{4} \arg \left( \sum_{n=0}^{N-1} \tilde{r}_n^4(\varepsilon) \right) \tag{A-8}$$

that again is a generalization of Eq. (9-16) to nonzero $\hat{\varepsilon}$, for the special case of $H = 4$.

For the ML estimates of symbol timing, we return to the exact forms in Eqs. (9-28) and (9-29). For example, for BPSK and a given value of $\theta_c$, differentiating Eq. (9-28) (with $\varepsilon$ replaced by $\hat{\varepsilon}$) with respect to $\hat{\varepsilon}$ and equating the result to zero gives the following:

$$\begin{aligned}
&\frac{d}{d\hat{\varepsilon}} \sum_{n=0}^{N-1} \ln\cosh \left[ \frac{\sqrt{2P_t}}{\sigma^2} \operatorname{Re}\left\{ \tilde{r}_n(\hat{\varepsilon}) e^{-j\theta_c} \right\} \right] \\
&\qquad = \sum_{n=0}^{N-1} \left( \tanh \left[ \frac{\sqrt{2P_t}}{\sigma^2} \operatorname{Re}\left\{ \tilde{r}_n(\hat{\varepsilon}) e^{-j\theta_c} \right\} \right] \right) \left( \frac{d}{d\hat{\varepsilon}} \left[ \frac{\sqrt{2P_t}}{\sigma^2} \operatorname{Re}\left\{ \tilde{r}_n(\hat{\varepsilon}) e^{-j\theta_c} \right\} \right] \right) \\
&\qquad = \sum_{n=0}^{N-1} \operatorname{Re}\left\{ \tilde{r}'_n(\hat{\varepsilon}) e^{-j\theta_c} \right\} \left( \tanh \left[ \frac{\sqrt{2P_t}}{\sigma^2} \operatorname{Re}\left\{ \tilde{r}_n(\hat{\varepsilon}) e^{-j\theta_c} \right\} \right] \right) = 0
\end{aligned} \tag{A-9}$$

where

$$\tilde{r}'_n(\hat{\varepsilon}) = \frac{2}{T}\int_{(n-\hat{\varepsilon})T}^{(n+1-\hat{\varepsilon})T} \tilde{r}(t)\,\frac{d}{d\hat{\varepsilon}}p(t-nT-\hat{\varepsilon}T)\,dt \tag{A-10}$$

$$= -\frac{2}{T}\int_{(n-\hat{\varepsilon})T}^{(n+1-\hat{\varepsilon})T} \tilde{r}(t)\,\frac{d}{dt}p(t-nT-\hat{\varepsilon}T)\,dt \tag{A-11}$$

Unfortunately Eq. (A-10) does not yield a closed-form solution for $\hat{\varepsilon}$. A similar situation takes place for QPSK, namely,

$$\frac{d}{d\hat{\varepsilon}}\sum_{n=0}^{N-1}\ln\left[\frac{1}{2}\cosh\left[\frac{\sqrt{2P_t}}{\sigma^2}\mathrm{Re}\left\{\tilde{r}_n(\hat{\varepsilon})e^{-j\theta_c}\right\}\right]+\frac{1}{2}\cosh\left[\frac{\sqrt{2P_t}}{\sigma^2}\mathrm{Im}\left\{\tilde{r}_n(\hat{\varepsilon})e^{-j\theta_c}\right\}\right]\right]$$

$$=\sum_{n=0}^{N-1}\frac{\sinh\left(\frac{\sqrt{2P_t}}{\sigma^2}\mathrm{Re}\left\{\tilde{r}_n(\hat{\varepsilon})e^{-j\theta_c}\right\}\right)\mathrm{Re}\left\{\tilde{r}'_n(\hat{\varepsilon})e^{-j\theta_c}\right\}+\sinh\left(\frac{\sqrt{2P_t}}{\sigma^2}\mathrm{Im}\left\{\tilde{r}_n(\hat{\varepsilon})e^{-j\theta_c}\right\}\right)\mathrm{Im}\left\{\tilde{r}'_n(\hat{\varepsilon})e^{-j\theta_c}\right\}}{\cosh\left[\frac{\sqrt{2P_t}}{\sigma^2}\mathrm{Re}\left\{\tilde{r}_n(\hat{\varepsilon})e^{-j\theta_c}\right\}\right]+\cosh\left[\frac{\sqrt{2P_t}}{\sigma^2}\mathrm{Im}\left\{\tilde{r}_n(\hat{\varepsilon})e^{-j\theta_c}\right\}\right]}$$

$$= 0 \tag{A-12}$$

If one now applies the approximations

$$\begin{aligned}\tanh x &\cong x\\ \sinh x &\cong x\\ \cosh x &\cong 1\end{aligned} \tag{A-13}$$

then Eqs. (A-9) and (A-12) simplify respectively to

$$\sum_{n=0}^{N-1}\mathrm{Re}\left\{\tilde{r}'_n(\hat{\varepsilon})\,e^{-j\theta_c}\right\}\mathrm{Re}\left\{\tilde{r}_n(\hat{\varepsilon})\,e^{-j\theta_c}\right\}$$

$$=\sum_{n=0}^{N-1}\left[\mathrm{Re}\left\{\tilde{r}'_n(\hat{\varepsilon})\left(\tilde{r}_n(\hat{\varepsilon})\right)^*\right\}+\mathrm{Re}\left\{\tilde{r}'_n(\hat{\varepsilon})\,\tilde{r}_n(\hat{\varepsilon})\,e^{-2j\theta_c}\right\}\right]=0 \tag{A-14}$$

and

$$\sum_{n=0}^{N-1}\left[\mathrm{Re}\left\{\tilde{r}_n\left(\hat{\varepsilon}\right)e^{-j\theta_c}\right\}\mathrm{Re}\left\{\tilde{r}'_n\left(\hat{\varepsilon}\right)e^{-j\theta_c}\right\}+\mathrm{Im}\left\{\tilde{r}_n\left(\hat{\varepsilon}\right)e^{-j\theta_c}\right\}\mathrm{Im}\left\{\tilde{r}'_n\left(\hat{\varepsilon}\right)e^{-j\theta_c}\right\}\right]$$

$$=\sum_{n=0}^{N-1}\mathrm{Re}\left\{\tilde{r}_n\left(\hat{\varepsilon}\right)\left(\tilde{r}'_n\left(\hat{\varepsilon}\right)\right)^*\right\}=0 \tag{A-15}$$

both of which require numerical solution for their respective ML estimates $\hat{\varepsilon}_2$ and $\hat{\varepsilon}_4$. Note from Eq. (A-15) that the ML estimate of symbol timing for the QPSK hypothesis is independent of the carrier phase estimate.

# Appendix 9-B

# ML Estimation of Carrier Phase for π/4-QPSK Modulation

To obtain the estimator of carrier phase needed for the GLRT involving $\pi/4$-QPSK modulation, we need to find the solution of

$$\hat{\theta}_{c,\pi/4-4}=\arg\max_{\theta_c}\left[\sum_{n=1,3,5,\cdots}^{N-1}\ln\left(\frac{1}{2}\left[\cosh\left[x_n\left(0;\theta_c\right)\right]+\cosh\left[x_n\left(1;\theta_c\right)\right]\right]\right)\right.$$

$$\left.+\sum_{n=0,2,4,\cdots}^{N-2}\ln\left(\frac{1}{2}\left[\cosh\left[y_n\left(0;\theta_c\right)\right]+\cosh\left[y_n\left(1;\theta_c\right)\right]\right]\right)\right] \tag{B-1}$$

with $x_n\left(q;\theta_c\right)$ and $y_n\left(q;\theta_c\right)$ as defined in Eq. (9-32). Applying the small argument approximations $\ln\left(1+x\right)\cong x$, $\cosh x\cong 1+x^2/2+x^4/24$ and ignoring second-order terms (since they contribute terms that do not depend on $\theta_c$) gives

$$
\begin{aligned}
&\hat{\theta}_{c,\pi/4-4} \\
&= \arg\max_{\theta_c} \left[ \frac{1}{48} \sum_{\substack{n=\\1,3,5,\cdots}}^{N-1} \left[ \left( \frac{\sqrt{2P_t}}{\sigma^2} \mathrm{Re}\left\{ \tilde{r}_n e^{-j\theta_c} \right\} \right)^4 + \left( \frac{\sqrt{2P_t}}{\sigma^2} \mathrm{Im}\left\{ \tilde{r}_n e^{-j\theta_c} \right\} \right)^4 \right] \right. \\
&\left. + \frac{1}{48} \sum_{\substack{n=\\0,2,4,\cdots}}^{N-2} \left[ \left( \frac{\sqrt{2P_t}}{\sigma^2} \mathrm{Re}\left\{ \tilde{r}_n e^{-j\pi/4} e^{-j\theta_c} \right\} \right)^4 + \left( \frac{\sqrt{2P_t}}{\sigma^2} \mathrm{Im}\left\{ \tilde{r}_n e^{-j\pi/4} e^{-j\theta_c} \right\} \right)^4 \right] \right] \\
&= \arg\max_{\theta_c} \left[ \frac{1}{48} \sum_{\substack{n=\\1,3,5,\cdots}}^{N-1} \left\{ \left[ \left( \frac{\sqrt{2P_t}}{\sigma^2} \mathrm{Re}\left\{ \tilde{r}_n e^{-j\theta_c} \right\} \right)^2 + \left( \frac{\sqrt{2P_t}}{\sigma^2} \mathrm{Im}\left\{ \tilde{r}_n e^{-j\theta_c} \right\} \right)^2 \right]^2 \right. \right. \\
&\left. -2 \left( \frac{\sqrt{2P_t}}{\sigma^2} \mathrm{Re}\left\{ \tilde{r}_n e^{-j\theta_c} \right\} \right)^2 \left( \frac{\sqrt{2P_t}}{\sigma^2} \mathrm{Im}\left\{ \tilde{r}_n e^{-j\theta_c} \right\} \right)^2 \right\} \\
&+ \frac{1}{48} \sum_{\substack{n=\\0,2,4,\cdots}}^{N-2} \left\{ \left[ \left( \frac{\sqrt{2P_t}}{\sigma^2} \mathrm{Re}\left\{ \tilde{r}_n e^{-j\pi/4} e^{-j\theta_c} \right\} \right)^2 + \left( \frac{\sqrt{2P_t}}{\sigma^2} \mathrm{Im}\left\{ \tilde{r}_n e^{-j\pi/4} e^{-j\theta_c} \right\} \right)^2 \right]^2 \right. \\
&\left. \left. -2 \left( \frac{\sqrt{2P_t}}{\sigma^2} \mathrm{Re}\left\{ \tilde{r}_n e^{-j\pi/4} e^{-j\theta_c} \right\} \right)^2 \left( \frac{\sqrt{2P_t}}{\sigma^2} \mathrm{Im}\left\{ \tilde{r}_n e^{-j\pi/4} e^{-j\theta_c} \right\} \right)^2 \right\} \right] \qquad \text{(B-2)}
\end{aligned}
$$

or equivalently

$$
\begin{aligned}
\hat{\theta}_{c,\pi/4-4} = \arg\min_{\theta_c} & \left[ \sum_{n=1,3,5,\cdots}^{N-1} \left( \mathrm{Re}\left\{ \tilde{r}_n e^{-j\theta_c} \right\} \right)^2 \left( \mathrm{Im}\left\{ \tilde{r}_n e^{-j\theta_c} \right\} \right)^2 \right. \\
& \left. + \sum_{n=0,2,4,\cdots}^{N-2} \left( \mathrm{Re}\left\{ \tilde{r}_n e^{-j\pi/4} e^{-j\theta_c} \right\} \right)^2 \left( \mathrm{Im}\left\{ \tilde{r}_n e^{-j\pi/4} e^{-j\theta_c} \right\} \right)^2 \right] \qquad \text{(B-3)}
\end{aligned}
$$

Letting $\tilde{r}_n = |\tilde{r}_n| e^{j\phi_n}$, the above becomes

$$\begin{aligned}
\hat{\theta}_{c,\pi/4-4} = \ & \arg\min_{\theta_c} \left[ \sum_{n=1,3,5,\cdots}^{N-1} |\tilde{r}_n|^4 \cos^2(\phi_n - \theta_c) \sin^2(\phi_n - \theta_c) \right. \\
& \left. + \sum_{n=0,2,4,\cdots}^{N-2} |\tilde{r}_n|^4 \cos^2\left(\phi_n - \frac{\pi}{4} - \theta_c\right) \sin^2\left(\phi_n - \frac{\pi}{4} - \theta_c\right) \right] \\
= \ & \arg\min_{\theta_c} \left[ \frac{1}{4} \sum_{n=1,3,5,\cdots}^{N-1} |\tilde{r}_n|^4 - \frac{1}{8} \sum_{n=1,3,5,\cdots}^{N-1} |\tilde{r}_n|^4 \cos\left[4(\phi_n - \theta_c)\right] \right. \\
& \left. - \frac{1}{8} \sum_{n=0,2,4,\cdots}^{N-2} |\tilde{r}_n|^4 \cos\left[4\left(\phi_n - \frac{\pi}{4} - \theta_c\right)\right] \right] \\
= \ & \arg\max_{\theta_c} \left[ \sum_{n=1,3,5,\cdots}^{N-1} |\tilde{r}_n|^4 \cos\left[4(\phi_n - \theta_c)\right] \right. \\
& \left. + \sum_{n=0,2,4,\cdots}^{N-2} |\tilde{r}_n|^4 \cos\left[4\left(\phi_n - \frac{\pi}{4} - \theta_c\right)\right] \right] \qquad \text{(B-4)}
\end{aligned}$$

This can also be written in the form

$$\begin{aligned}
& \hat{\theta}_{c,\pi/4-4} \\
& = \arg\max_{\theta_c} \left[ \mathrm{Re}\left\{ \sum_{n=1,3,5,\cdots}^{N-1} \tilde{r}_n^4 \right\} \mathrm{Re}\left\{e^{j4\theta_c}\right\} + \mathrm{Im}\left\{ \sum_{n=1,3,5,\cdots}^{N-1} \tilde{r}_n^4 \right\} \mathrm{Im}\left\{e^{j4\theta_c}\right\} \right. \\
& \quad \left. + \mathrm{Re}\left\{ \sum_{n=0,2,4,\cdots}^{N-2} \tilde{r}_n^4 \right\} \mathrm{Re}\left\{e^{j4\left(\frac{\pi}{4}+\theta_c\right)}\right\} + \mathrm{Im}\left\{ \sum_{n=0,2,4,\cdots}^{N-2} \tilde{r}_n^4 \right\} \mathrm{Im}\left\{e^{j4\left(\frac{\pi}{4}+\theta_c\right)}\right\} \right] \\
& = \arg\max_{\theta_c} \left\{ \left| \sum_{n=1,3,5,\cdots}^{N-1} \tilde{r}_n^4 \right| \cos\left[ \arg\left( \sum_{n=1,3,5,\cdots}^{N-1} \tilde{r}_n^4 \right) - 4\theta_c \right] \right. \\
& \quad \left. + \left| \sum_{n=0,2,4,\cdots}^{N-2} \tilde{r}_n^4 \right| \cos\left[ \arg\left( \sum_{n=0,2,4,\cdots}^{N-2} \tilde{r}_n^4 \right) - \pi - 4\theta_c \right] \right\} \\
& = \arg\max_{\theta_c} \left\{ \left| \sum_{n=1,3,5,\cdots}^{N-1} \tilde{r}_n^4 \right| \cos\left[ \arg\left( \sum_{n=1,3,5,\cdots}^{N-1} \tilde{r}_n^4 \right) - 4\theta_c \right] \right. \\
& \quad \left. - \left| \sum_{n=0,2,4,\cdots}^{N-2} \tilde{r}_n^4 \right| \cos\left[ \arg\left( \sum_{n=0,2,4,\cdots}^{N-2} \tilde{r}_n^4 \right) - 4\theta_c \right] \right\} \qquad \text{(B-5)}
\end{aligned}$$

However, using the trigonometric identity

$$A\cos(a-x) - B\cos(b-x) = (A\cos a - B\cos b)\cos x + (A\sin a - B\sin b)\sin x$$

$$= \sqrt{(A\cos a - B\cos b)^2 + (A\sin a - B\sin b)^2}\cos(x-\eta) \tag{B-6}$$

$$\eta = \tan^{-1}\frac{A\sin a - B\sin b}{A\cos a - B\cos b}$$

in Eq. (B-5) gives

$$\hat{\theta}_{c,\pi/4-4} = \arg\max_{\theta_c}\left[\cos(4\theta_c - \eta)\right] \tag{B-7}$$

where

$$\eta = \tan^{-1}\frac{\operatorname{Im}\left\{\sum_{n=1,3,5,\cdots}^{N-1}\tilde{r}_n^4\right\} - \operatorname{Im}\left\{\sum_{n=0,2,4,\cdots}^{N-2}\tilde{r}_n^4\right\}}{\operatorname{Re}\left\{\sum_{n=1,3,5,\cdots}^{N-1}\tilde{r}_n^4\right\} - \operatorname{Re}\left\{\sum_{n=0,2,4,\cdots}^{N-2}\tilde{r}_n^4\right\}} \tag{B-8}$$

Finally then,

$$\hat{\theta}_{c,\pi/4-4} = \frac{1}{4}\tan^{-1}\frac{\operatorname{Im}\left\{\sum_{n=1,3,5,\cdots}^{N-1}\tilde{r}_n^4\right\} - \operatorname{Im}\left\{\sum_{n=0,2,4,\cdots}^{N-2}\tilde{r}_n^4\right\}}{\operatorname{Re}\left\{\sum_{n=1,3,5,\cdots}^{N-1}\tilde{r}_n^4\right\} - \operatorname{Re}\left\{\sum_{n=0,2,4,\cdots}^{N-2}\tilde{r}_n^4\right\}} \tag{B-9}$$

$$= \frac{1}{4}\tan^{-1}\frac{\operatorname{Im}\left\{\sum_{n=0}^{N-1}(-1)^{n-1}\tilde{r}_n^4\right\}}{\operatorname{Re}\left\{\sum_{n=0}^{N-1}(-1)^{n-1}\tilde{r}_n^4\right\}} = \frac{1}{4}\arg\left(\sum_{n=0}^{N-1}(-1)^{n-1}\tilde{r}_n^4\right) \tag{B-10}$$

$$= \frac{1}{4}\arg\left(\sum_{n=0}^{N-1}\left(\tilde{r}_n e^{j(\pi/4)(n-1)}\right)^4\right) = \frac{1}{4}\arg\left(\sum_{n=0}^{N-1}\left(\tilde{r}_n e^{-j(\pi/4)I_n}\right)^4\right)$$

where $I_n$ is the indicator variable defined by

$$I_n = \begin{cases} 0, & n \text{ odd} \\ 1, & n \text{ even} \end{cases} \tag{B-11}$$